

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: February 19, 2004, 13:46:37 ; Search time 14114 Seconds
(without alignments)
17476.634 Million cell updates/sec

Title: US-09-964-956-12
Perfect score: 5691
Sequence: 1 atgaagccatgcctcgaa.....gcttagacagctgaataaa 5691

Scoring table: IDENTITY NUC
Gapop 10.0, Gapext 1.0

Searched: 3470272 seqs, 21671516995 residues
Total number of hits satisfying chosen parameters: 6940544

Minimum DB seq length: 0
Maximum DB seq length: 200000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :
1: gb_ba:*
2: gb_hgt:*
3: gb_in:*
4: gb_om:*
5: gb_ov:*
6: gb_pat:*
7: gb_ph:*
8: gb_pl:*
9: gb_pr:*
10: gb_ro:*
11: gb_sts:*
12: gb_sy:*
13: gb_un:*
14: gb_vl:*
15: em_ba:*
16: em_fun:*
17: em_hum:*
18: em_in:*
19: em_mu:*
20: em_om:*
21: em_or:*
22: em_ov:*
23: em_pat:*
24: em_ph:*
25: em_pl:*
26: em_ro:*
27: em_sts:*
28: em_un:*
29: em_vl:*
30: em_hgt_hum:*
31: em_hgt_in:*
32: em_hgt_inv:*
33: em_hgt_other:*
34: em_hgt_mus:*
35: em_hgt_pln:*
36: em_hgt_rod:*
37: em_hgt_mam:*
38: em_hgt_vrt:*
39: em_ey:*
40: em_hgtgo_hum:*
41: em_hgtgo_mus:*
42: em_hgtgo_other:*

score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	5691	100.0	5691	6	AX503782
2	5680.6	99.8	6367	6	AX42982 Sequence
3	4714.2	82.8	7022	10	AB073228
4	2939.6	51.7	6387	6	AX405895
5	2938	51.6	6263	9	AB007932
6	2778.4	48.8	6730	6	AX305728
7	2778.4	48.8	6730	10	D86949
8	2659.4	46.7	6147	6	AX704766
9	2578.4	45.3	6466	10	D86948
10	2535.4	44.6	6039	9	HSSEXGENE
11	2533.8	44.5	5895	6	AX686475
12	2423.2	42.6	5263	9	HSNOVPROT
13	2387.8	42.0	6893	10	AK122289
14	2366.8	41.6	6163	5	XELPLEX
15	2312.8	40.6	5982	10	D86950
16	2240.2	39.4	5986	9	AB046770
17	2146.8	37.7	3985	9	HSOCTPROT
18	2015.6	35.4	6693	6	BD183298
19	1554	27.3	2875	9	BC009343
20	1377	24.2	2015	9	BC028744
21	1371.8	24.1	3310	9	AK123428
22	1188.4	20.9	179145	9	AC009785
23	1185.2	20.8	188090	9	AC011625
24	1177.4	20.7	179592	2	AC009723
25	1167.4	20.5	2284	10	BC056475
26	1157.8	20.3	3666	6	AX833678
27	1157.8	20.3	3666	9	AK095606
28	1026.6	18.0	3556	9	HSMB05617
29	930.8	16.4	221370	2	AC126205
30	930.8	16.4	228898	2	AC131554
31	930.8	16.4	229922	2	AC102141
32	915.2	16.1	1817	9	BC006193
33	878.2	15.4	4445	9	AK126101
34	793.8	13.9	2597	6	AX574564
35	793.8	13.9	2597	9	AX358850
36	759.8	13.4	1029	9	HSMB02037
37	673.6	11.8	776	4	AF176418
38	630.8	11.1	3560	6	AX747045
39	630.8	11.1	3560	9	AK091538
40	625	11.0	74655	2	AC025595
41	613.2	10.8	1403	10	BC024509
42	598.4	10.5	56201	2	AC144970
43	591.4	10.4	601	6	AX088116
44	586.6	10.3	2414	6	BD127385
45	586.6	10.3	2414	9	AK074960

ALIGNMENTS

RESULT 1	AX503782	5691 bp	DNA	linear	PAT 27-SEP-2002
LOCUS	AX503782				
DEFINITION	Sequence 12 from Patent WO0226826.				
ACCESSION	AX503782				
VERSION	AX503782.1	GI:23385966			
KEYWORDS					
SOURCE					
ORGANISM	Homo sapiens (human)				
REFERENCE					
AUTHORS	Gerlach, V.L., MacDougall, J.R., Smithson, G., Miller, I., Stone, D., Gunter, E., Ellemann, K., Grose, W.M., Alsbrook, J.P., Lepley, D.M., Burgess, C.E., Padigara, M., Kekuda, R., Szytek, K.A., Leach, M.D. and				

Shinkets, R.A.
Proteins and nucleic acids encoding same
Patent: WO 026826-A 12 04-APR-2002;
JOURNAL Curagen Corporation (US)

FEATURES

Source

1. 5691
/organism="Homo sapiens"
/mol_type="unassigned DNA"
/db_xref="taxon:9606"

ORIGIN

Query Match 100.0%; Score 5691; DB 6; Length 5691;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 5691; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

1 ATGAAAGCCATGCCCCCTGAACTGACCTGCTTCTCTCCCACTCTCTCATGTGGCATG 60
1 ATGAAAGCCATGCCCCCTGAACTGACCTGCTTCTCTCCCACTCTCTCATGTGGCATG 60
61 GGCTCTTCCACTTGTCTCACTCCGCGACGCCCCCGTGTCCAGAAAGCAGCGGTATT 120
61 GGCTCTTCCACTTGTCTCACTCCGCGACGCCCCCGTGTCCAGAAAGCAGCGGTATT 120
121 GTCACTTCCGAGAGAGACCCGCGAGGGTTTCAATCACTGTGTGATGAGAGACA 180
121 GTCACTTCCGAGAGAGACCCGCGAGGGTTTCAATCACTGTGTGATGAGAGACA 180
181 GGACACATTACTTGGGGGCGCTCAATGCGATTTCAGAGCTTTCAGAGCACTGAAAGTTC 240
181 GGACACATTACTTGGGGGCGCTCAATGCGATTTCAGAGCTTTCAGAGCACTGAAAGTTC 240
241 TTGGTGAAGCATGAGACAGGGCCGGAAGACCAACCCCAAGTTTACCACCCCGCATC 300
241 TTGGTGAAGCATGAGACAGGGCCGGAAGACCAACCCCAAGTTTACCACCCCGCATC 300
301 GTCCAGACCTGCAATGAGCCCTTGACCAACCAATGTCAACAAAGATGCTTCATTA 360
301 GTCCAGACCTGCAATGAGCCCTTGACCAACCAATGTCAACAAAGATGCTTCATTA 360
361 GACTCAAGAGAGAAAGGCTGATTTGCTGTGGAGCCCTGTCACAGGCACTTCAGACTG 420
361 GACTCAAGAGAGAAAGGCTGATTTGCTGTGGAGCCCTGTCACAGGCACTTCAGACTG 420
421 CTGAGGCTGAGAGACCTTTCAGCTGGGGAGCCCTTATCATTAAGAGAGCACTACTG 480
421 CTGAGGCTGAGAGACCTTTCAGCTGGGGAGCCCTTATCATTAAGAGAGCACTACTG 480
481 TCAGGTTCACAGAGAGGCGCTCAGTCTTTGGAGTGAATCGTCTCTACAGAACCTGGAT 540
481 TCAGGTTCACAGAGAGGCGCTCAGTCTTTGGAGTGAATCGTCTCTACAGAACCTGGAT 540
541 GACAAGCTGTTCAATTCACAGCGCAGTGTAGTGGAAAGCCGAGTATTTTCCACCATCTCC 600
541 GACAAGCTGTTCAATTCACAGCGCAGTGTAGTGGAAAGCCGAGTATTTTCCACCATCTCC 600
601 AGCCGGAAGTGAACCAAGACTCTGAGAGCGAGTGGATGTTGGGATACGCTTCCATGAT 660
601 AGCCGGAAGTGAACCAAGACTCTGAGAGCGAGTGGATGTTGGGATACGCTTCCATGAT 660
661 GAGTTCGTGGCTCGATGATTAGATCCCTTCGAGCACTTCACCATCATCCCTGACTTT 720
661 GAGTTCGTGGCTCGATGATTAGATCCCTTCGAGCACTTCACCATCATCCCTGACTTT 720
721 GATATCTACTATGTCTATGTGTTTGAAGTGGCACTTGTCTACTTTTGAACCTTCAA 780
721 GATATCTACTATGTCTATGTGTTTGAAGTGGCACTTGTCTACTTTTGAACCTTCAA 780
781 CCTGAGATGTGTCTCCACAGGCTCCACACCAAGAGAGCAGGTGTATATCAATCAAGTTC 840
781 CCTGAGATGTGTCTCCACAGGCTCCACACCAAGAGAGCAGGTGTATATCAATCAAGTTC 840
841 GTGAGGCTTTGCAAGAGAGACAGGCTTCAACTCTTATGTAGAGTCCCATGAGCTGT 900

```

```

841 GTGAGGCTTTGCAAGAGAGACAGGCTTCAACTCTTATGTAGAGTCCCATGAGCTGT 900
901 GAGGCGAGTGGGGTGAAGTACCGCTCTGTCAGAGCTCTCTACTGTCTCAAAAGCGGGGCC 960
901 GAGGCGAGTGGGGTGAAGTACCGCTCTGTCAGAGCTCTCTACTGTCTCAAAAGCGGGGCC 960
961 GTGCTGGCAGAGACCTCTGAGTCCATCCAGATGATACCTGTCTTCAACGCTTCTCC 1020
961 GTGCTGGCAGAGACCTCTGAGTCCATCCAGATGATACCTGTCTTCAACGCTTCTCC 1020
1021 AAGGCCAGAAAGCGAAATGAATTCCTGATGATGCGCCCTGTGTCATCTTATCTTG 1080
1021 AAGGCCAGAAAGCGAAATGAATTCCTGATGATGCGCCCTGTGTCATCTTATCTTG 1080
1081 AAGCAGATTAATGACCGCATTAAGAGCGGTGAGCTTGTATACCGGGCGAGGACG 1140
1081 AAGCAGATTAATGACCGCATTAAGAGCGGTGAGCTTGTATACCGGGCGAGGACG 1140
1141 CTGAGCTGGCTGGCTCAAGGTGAAGACATCCCTGACAGAGTGGCTTTAACATTT 1200
1141 CTGAGCTGGCTGGCTCAAGGTGAAGACATCCCTGACAGAGTGGCTTTAACATTT 1200
1201 GACGATTACTTCTGTGGCTTGAAGATGAATGCTCCCTGGAGAGTCCGACATGTTGGT 1260
1201 GACGATTACTTCTGTGGCTTGAAGATGAATGCTCCCTGGAGAGTCCGACATGTTGGT 1260
1261 GGAATTCCCGCTTCTCAAGAGAGACAGGACCCGATGACGTCTGTATGATGTCTAC 1320
1261 GGAATTCCCGCTTCTCAAGAGAGACAGGACCCGATGACGTCTGTATGATGTCTAC 1320
1321 AAGAACACTCTCTGGCTTTGTGGGACCAAAAGTGGCAGCTGAAGAAATCCGGGTG 1380
1321 AAGAACACTCTCTGGCTTTGTGGGACCAAAAGTGGCAGCTGAAGAAATCCGGGTG 1380
1381 GATGAGCCAGAGGCAAGCCCTCAATGATGACGAGTGTGAGAGTGTGAGCCGCCCA 1440
1381 GATGAGCCAGAGGCAAGCCCTCAATGATGACGAGTGTGAGAGTGTGAGCCGCCCA 1440
1441 GTCTCCGAGATATGCTCTTCTCCAGAGACCAAGCACTCTATCATGTCAAGAGAG 1500
1441 GTCTCCGAGATATGCTCTTCTCCAGAGACCAAGCACTCTATCATGTCAAGAGAG 1500
1501 CAGCTCAACAGAGTCCCTGTGAGTCTGTGGTCAATGATGAGCTGGCGAGTGCCTT 1560
1501 CAGCTCAACAGAGTCCCTGTGAGTCTGTGGTCAATGATGAGCTGGCGAGTGCCTT 1560
1561 GGCTCAAGCGACCCCACTGTGCTGTGTGTCTGCAACAAGTGCACCCGGAAGAG 1620
1561 GGCTCAAGCGACCCCACTGTGCTGTGTGTCTGCAACAAGTGCACCCGGAAGAG 1620
1621 GAGTCAAGGCAACCCCACTGTGCTGTGTGTCTGCAACAAGTGCACCCGGAAGAG 1680
1621 GAGTCAAGGCAACCCCACTGTGCTGTGTGTCTGCAACAAGTGCACCCGGAAGAG 1680
1681 CTGACGCTCATCCCAACATATCTCCGTCTCTCAATGACAGTGTGTCTGTCTGAG 1740
1681 CTGACGCTCATCCCAACATATCTCCGTCTCTCAATGACAGTGTGTCTGTCTGAG 1740
1741 AGCTAATATGTCCTGGAGCTGACAGTGGCTCAACTGCACTTTGAGAGACTGTCAAG 1800
1741 AGCTAATATGTCCTGGAGCTGACAGTGGCTCAACTGCACTTTGAGAGACTGTCAAG 1800
1801 ATGATGAGCTGGTCTGTGGCAATCAAGATCCAGTGTATCTCCCTGACACCAAGAGGTG 1860
1801 ATGATGAGCTGGTCTGTGGCAATCAAGATCCAGTGTATCTCCCTGACACCAAGAGGTG 1860
1861 CCGCGATATCAAGAGATGGGAGACCAATGTCTGACCTTCAGCTCAATCAAG 1920
1861 CCGCGATATCAAGAGATGGGAGACCAATGTCTGACCTTCAGCTCAATCAAG 1920
1921 GAGACCGGATGACCTTTCAGACACAGCTTGTCTTCAATTTGACAGCTTCAAT 1980
1921 GAGACCGGATGACCTTTCAGACACAGCTTGTCTTCAATTTGACAGCTTCAAT 1980

```

QY 1981 TCGTGGCTGTCTCTCGGTGAGAGTCAATACCGCTGCCAATGCTGTAATATACCGGCAATGTC 2040
Db 1981 TCGTGGCTGTCTCTCGGTGAGAGTCAATACCGCTGCCAATGCTGTAATATACCGGCAATGTC 2040
QY 2041 TGCACCCATGACCCCAAGACCTGCTCTTCCAGAGAGGCGAGTGAAGCTGCCAGAGAC 2100
Db 2041 TGCACCCATGACCCCAAGACCTGCTCTTCCAGAGAGGCGAGTGAAGCTGCCAGAGAC 2100
QY 2101 TGGCCCCAGCTGCTGCGAGTGAACAAGATCTGTGCTCCCTGGAGAGTGAATCAAGCCATAC 2160
Db 2101 TGGCCCCAGCTGCTGCGAGTGAACAAGATCTGTGCTCCCTGGAGAGTGAATCAAGCCATAC 2160
QY 2161 ACGCTGAAGGCCAAGAACCTCCCGCAGGCCAGTCTGTGGGCAAGCTGTGCTAAGATGATC 2220
Db 2161 ACGCTGAAGGCCAAGAACCTCCCGCAGGCCAGTCTGTGGGCAAGCTGTGCTAAGATGATC 2220
QY 2221 CTCAAATTCAGGGCAGGAGCAGAGAGTGCCTCCGCTTCAACAGCTCCAGGCTA 2280
Db 2221 CTCAAATTCAGGGCAGGAGCAGAGAGTGCCTCCGCTTCAACAGCTCCAGGCTA 2280
QY 2281 CAGTGGCCAGAACACCTCTTATTCCTATGAGAGATGAGATCAACAACCTGCGGTGAG 2340
Db 2281 CAGTGGCCAGAACACCTCTTATTCCTATGAGAGATGAGATCAACAACCTGCGGTGAG 2340
QY 2341 TTGACAGTCTGTGAGATGGGCACTTCAACATGACAAACCGAGCTCAAGATTAAGTTTAC 2400
Db 2341 TTGACAGTCTGTGAGATGGGCACTTCAACATGACAAACCGAGCTCAAGATTAAGTTTAC 2400
QY 2401 CTCTAACAGTGTGAGAGCCTATGCTGAGAGCTGCGGCTGCTCAAGGCTGACCCGAGAC 2460
Db 2401 CTCTAACAGTGTGAGAGCCTATGCTGAGAGCTGCGGCTGCTCAAGGCTGACCCGAGAC 2460
QY 2461 TTGCGATGTGGCTGGTGGCCAGAGGCCCAAGCTGACCTCTGCGCCAGCACTGCGCTGCC 2520
Db 2461 TTGCGATGTGGCTGGTGGCCAGAGGCCCAAGCTGACCTCTGCGCCAGCACTGCGCTGCC 2520
QY 2521 CAGGAGAGCCAGTGTGCTGAGAGCTGTCTGTGTGCTGCAAAAGCAAGTGCACAAACCCCGGATC 2580
Db 2521 CAGGAGAGCCAGTGTGCTGAGAGCTGTCTGTGTGCTGCAAAAGCAAGTGCACAAACCCCGGATC 2580
QY 2581 ACAGAGATTAATCCCGGTGACAGAGCCCGGAGAGGGGAGCCAAAGTCACTATCCAGAGG 2640
Db 2581 ACAGAGATTAATCCCGGTGACAGAGCCCGGAGAGGGGAGCCAAAGTCACTATCCAGAGG 2640
QY 2641 GAGAACTGTGGGCTGGAAATTTGCGCAATCGCTCTCCATGTCAAGAGTTGCTGGGCTGAG 2700
Db 2641 GAGAACTGTGGGCTGGAAATTTGCGCAATCGCTCTCCATGTCAAGAGTTGCTGGGCTGAG 2700
QY 2701 TGCACCCCTTTATGAGATGTTTATCATCCCTGAGAGAGATGATGATGAGATGAGAGG 2760
Db 2701 TGCACCCCTTTATGAGATGTTTATCATCCCTGAGAGAGATGATGATGAGATGAGAGG 2760
QY 2761 GCCAAGCCCAAGCAGATGCAAGGCTTCGTGAGATCTGCTGTGTGTGTGTGTGTGAA 2820
Db 2761 GCCAAGCCCAAGCAGATGCAAGGCTTCGTGAGATCTGCTGTGTGTGTGTGTGTGAA 2820
QY 2821 TTTCATAGCCCGGTCTCTCAAGCTTAATTAATTATGACATGACATGACATGACATGAA 2880
Db 2821 TTTCATAGCCCGGTCTCTCAAGCTTAATTAATTATGACATGACATGACATGACATGAA 2880
QY 2881 CCCAGCCGGGGCCCAATGTCGCGAGAGAGCCCAAGTGAACATCAAGGACCAACCTGAT 2940
Db 2881 CCCAGCCGGGGCCCAATGTCGCGAGAGAGCCCAAGTGAACATCAAGGACCAACCTGAT 2940
QY 2941 GCCGGAAGCAAGTGTGTGTATGTTTGAAGAGAGCCCTGTCTTTCACAGGAGATCT 3000
Db 2941 GCCGGAAGCAAGTGTGTGTATGTTTGAAGAGAGCCCTGTCTTTCACAGGAGATCT 3000
QY 3001 CCATCTTACATTTGTCTGCAACCAACATCTCAGATGAGGTGCTGAGATGAAGGTGTG 3060
Db 3001 CCATCTTACATTTGTCTGCAACCAACATCTCAGATGAGGTGCTGAGATGAAGGTGTG 3060

QY 3061 GTGCAGGTGACAGAGGCCAAGATCCACAGGACCTGTGTTCAATATGTGAGAACCC 3120
Db 3061 GTGCAGGTGACAGAGGCCAAGATCCACAGGACCTGTGTTCAATATGTGAGAACCC 3120
QY 3121 ACCATCTGTGGATTTGAGATCCAGAAATGAGAGCAATGTCAATGAGAAACACCCATGCCGTA 3180
Db 3121 ACCATCTGTGGATTTGAGATCCAGAAATGAGAGCAATGTCAATGAGAAACACCCATGCCGTA 3180
QY 3181 TGGGGAGCCCACTGACCTCATACAGAACCCCAAGATCCGTGCCAATGAGAGGAG 3240
Db 3181 TGGGGAGCCCACTGACCTCATACAGAACCCCAAGATCCGTGCCAATGAGAGGAG 3240
QY 3241 GAGCATATCAATATCTGAGAGTTCTGAACGTTATGAGATGACCTGTCAAGGCGCGGC 3300
Db 3241 GAGCATATCAATATCTGAGAGTTCTGAACGTTATGAGATGACCTGTCAAGGCGCGGC 3300
QY 3301 CTGCTGTGGTCTGACCAAGCAGTCAAGCTGACCCGAGAGGCCAGAGATTGGCTTC 3360
Db 3301 CTGCTGTGGTCTGACCAAGCAGTCAAGCTGACCCGAGAGGCCAGAGATTGGCTTC 3360
QY 3361 ATCTGAGACAGTCCAGTCTCTGCTCATCTCTCAACAGACCTTACCTAATATCC 3420
Db 3361 ATCTGAGACAGTCCAGTCTCTGCTCATCTCTCAACAGACCTTACCTAATATCC 3420
QY 3421 AACCCGGTGTGAGGCTTTGGTCCCTGAGGATCTGAGAGCTCAAGCCTGAGAGCC 3480
Db 3421 AACCCGGTGTGAGGCTTTGGTCCCTGAGGATCTGAGAGCTCAAGCCTGAGAGCC 3480
QY 3481 ATCATCTTAAAGGCAAGAACCTGATCCCGCTGTGCTGGGGCAACCTGAAGCTGAAC 3540
Db 3481 ATCATCTTAAAGGCAAGAACCTGATCCCGCTGTGCTGGGGCAACCTGAAGCTGAAC 3540
QY 3541 TACATCTGTGCTGTGGGAGAGAGCCGTCACACGTCACCGTGTGATGCTCCAGCTGTC 3600
Db 3541 TACATCTGTGCTGTGGGAGAGAGCCGTCACACGTCACCGTGTGATGCTCCAGCTGTC 3600
QY 3601 TGGGATCCCCCAACCTCATGCGCAGGCAAAAGTATGAGCCGTGTGCTGAGATGAG 3660
Db 3601 TGGGATCCCCCAACCTCATGCGCAGGCAAAAGTATGAGCCGTGTGCTGAGATGAG 3660
QY 3661 TACTCCCCGGGATGTGTATATGTCGCCGAGCAAGCCGCTAGCCTGCGGCAATGATC 3720
Db 3661 TACTCCCCGGGATGTGTATATGTCGCCGAGCAAGCCGCTAGCCTGCGGCAATGATC 3720
QY 3721 AGCATGCAAGTGTGGGAGCTCTCATATTTTCAATGATGATGATGATGATGATGATG 3780
Db 3721 AGCATGCAAGTGTGGGAGCTCTCATATTTTCAATGATGATGATGATGATGATGATG 3780
QY 3781 AAACGCAAGTCCCGCAAAAGTGAACCTCAAGCTGAGAGGCTGACATGACATGACATG 3840
Db 3781 AAACGCAAGTCCCGCAAAAGTGAACCTCAAGCTGAGAGGCTGACATGACATGACATG 3840
QY 3841 CTGAGATCCGTTGTGGCCCTGAGAGTCAAGAGACCTTTTGGAGCTGAGAGCGACATC 3900
Db 3841 CTGAGATCCGTTGTGGCCCTGAGAGTCAAGAGACCTTTTGGAGCTGAGAGCGACATC 3900
QY 3901 CATGAGCTGACAGTGAACCTGATGAGAGCCGGGATTCCTGCTGAGACTATTAACCTTAC 3960
Db 3901 CATGAGCTGACAGTGAACCTGATGAGAGCCGGGATTCCTGCTGAGACTATTAACCTTAC 3960
QY 3961 ACCATGGGGTGTGTTCCAGAGATTTGAAGACCAACCTGTCTCCGAGACCTTGAAGTC 4020
Db 3961 ACCATGGGGTGTGTTCCAGAGATTTGAAGACCAACCTGTCTCCGAGACCTTGAAGTC 4020
QY 4021 CCGGGCTACCGGAGAGAGCTGTGAGAAAGCTTAAAGCTCTTCCGCCAGCTCATAC 4080
Db 4021 CCGGGCTACCGGAGAGAGCTGTGAGAAAGCTTAAAGCTCTTCCGCCAGCTCATAC 4080
QY 4081 AACAAAGTGTCTGTGCTGCTTCAATCCGAGAGCTTGAAGTCCAGAGCTTACCTTCCATG 4140
Db 4081 AACAAAGTGTCTGTGCTGCTTCAATCCGAGAGCTTGAAGTCCAGAGCTTACCTTCCATG 4140
QY 4141 CGGACCGTGGCAAGCTGGCTCATCATCATGACCGGTGCTGAGAGCAAGCTGAGATAC 4200

Db	4141	CGGACCGTGGCAAGTGGCCCTCACTCATATGACCGTGTGGAGCAAGCTGGAGTAC	4200
Qy	4201	GGCAGTGAATGCTGAAGCAAGCTGTGTGGCCGACCTCATTTGAACAAGAACTTGGAGACGAG	4260
Db	4201	GGCAGTGAATGCTGAAGCAAGCTGTGTGGCCGACCTCATTTGAACAAGAACTTGGAGAGCAAG	4260
Qy	4261	AACCACTTAAGCTGTGTCTGACGAGAGCTGAAGTCAAGTGTGCTGAGAGATGCTGACCAAT	4320
Db	4261	AACCACTTAAGCTGTGTCTGACGAGAGCTGAAGTCAAGTGTGCTGAGAGATGCTGACCAAT	4320
Qy	4321	TGGTTTAACTTTCCCTCCTCAACAAGTTCCTCAAGAAAGTGTCTGGGAGGCCCTCTTCTCC	4380
Db	4321	TGGTTTAACTTTCCCTCCTCAACAAGTTCCTCAAGAAAGTGTCTGGGAGGCCCTCTTCTCC	4380
Qy	4381	CTGTTCTGTGCCATCAAGCAGATGTGAGAAAGGGCCCAATTGACCCCATCAAGGGCGAG	4440
Db	4381	CTGTTCTGTGCCATCAAGCAGATGTGAGAAAGGGCCCAATTGACCCCATCAAGGGCGAG	4440
Qy	4441	GGCCGCTACTCTCTTGAAGGAGACAAAGTCAATCCGACAGATTTGACTTACAAAAACCTTG	4500
Db	4441	GGCCGCTACTCTCTTGAAGGAGACAAAGTCAATCCGACAGATTTGACTTACAAAAACCTTG	4500
Qy	4501	GTCTGAGCTGTGTGACGCCACATTCGACAGCCCGAGGTCCCAAGTAAAGATCTCTC	4560
Db	4501	GTCTGAGCTGTGTGACGCCACATTCGACAGCCCGAGGTCCCAAGTAAAGATCTCTC	4560
Qy	4561	AAGTGTGAACACATCACTCAAGTCAAGAGAAAGTTGTGGATATCCATCTTCAAGAAATGTG	4620
Db	4561	AAGTGTGAACACATCACTCAAGTCAAGAGAAAGTTGTGGATATCCATCTTCAAGAAATGTG	4620
Qy	4621	CCTTGTCTCCACCGGCCCAAGCTGTGACAGATATGTGATCTGAGTGTGGCGACAAAGAAATGGG	4680
Db	4621	CCTTGTCTCCACCGGCCCAAGCTGTGACAGATATGTGATCTGAGTGTGGCGACAAAGAAATGGG	4680
Qy	4681	GCAAGGATGATCTTGTGAAGATGAAAGACATCAACCAACGAAATTGAGATATTTGGAAGGGA	4740
Db	4681	GCAAGGATGATCTTGTGAAGATGAAAGACATCAACCAACGAAATTGAGATATTTGGAAGGGA	4740
Qy	4741	CTGAACACACTGGCCCACTACAGGTGTGCAGATGGTTCCTGTGTGGCATTTAGTGTCAAG	4800
Db	4741	CTGAACACACTGGCCCACTACAGGTGTGCAGATGGTTCCTGTGTGGCATTTAGTGTCAAG	4800
Qy	4801	CAGGTGAACGCTATTAAGAGAGTGAACAATCCACCGTCTCCAGACCTCAAGAGTAA	4860
Db	4801	CAGGTGAACGCTATTAAGAGAGTGAACAATCCACCGTCTCCAGACCTCAAGAGTAA	4860
Qy	4861	TATGAAAACATGATCCGGTACACGGGCAGCCCCGACAGCCTCCGCTCAAGACACTATG	4920
Db	4861	TATGAAAACATGATCCGGTACACGGGCAGCCCCGACAGCCTCCGCTCAAGACACTATG	4920
Qy	4921	ATCACTCCGACCTGTGAAGTGTGAAGTCAAGATGTGTGGACCTAATGTAAAGAACACAGAGAC	4980
Db	4921	ATCACTCCGACCTGTGAAGTGTGAAGTCAAGATGTGTGGACCTAATGTAAAGAACACAGAGAC	4980
Qy	4981	GAGACCAAGAGAGGGGAGACCGGGAGAGCAAGATGTGTCTGAAATCTTAACTGACCGGA	5040
Db	4981	GAGACCAAGAGAGGGGAGACCGGGAGAGCAAGATGTGTCTGAAATCTTAACTGACCGGA	5040
Qy	5041	CTCTGTGGCACTTAAGGCACTGTGCAAGAGTTTGTGATGACCTCTTTAGACCAATCTTC	5100
Db	5041	CTCTGTGGCACTTAAGGCACTGTGCAAGAGTTTGTGATGACCTCTTTAGACCAATCTTC	5100
Qy	5101	AGCAGCGGCACACCGTGTGTGTGCTGTGCTGTGCTGTGACATCAAGTACATGTTTACTTCTG	5160
Db	5101	AGCAGCGGCACACCGTGTGTGTGCTGTGCTGTGCTGTGACATCAAGTACATGTTTACTTCTG	5160
Qy	5161	GATGACAGAGCTGATTAACATGTGCAATCCCGACAGTCCGCAATCTCGGAAGAGGC	5220
Db	5161	GATGACAGAGCTGATTAACATGTGCAATCCCGACAGTCCGCAATCTCGGAAGAGGC	5220
Qy	5221	AATTGCTGTGCTGTGAGTTTGTGGTCAACATGATCAAGAACCCGCAATTTGTGTTGAC	5280

Db	5221	AATTGCTGGCCCTGAGAGTTTGGGTCAACATGATCAAGAACCCGAGTTGTGTGAC	5285
QY	5281	ATTCATATAGAACACATATCAAGAGCCCTGCTCTCTGTGTGGTCACTTCATGAGAC	5340
Db	5281	ATTCATATAGAACACATATCAAGAGCCCTGCTCTCTGTGTGGTCACTTCATGAGAC	5340
QY	5341	TCTTGTCTCCAGTCAAGCAACCCGCTGGGCAAGGACTCGCCCTCCAAACAGCTGTGTAT	5400
Db	5341	TCTTGTCTCCAGTCAAGCAACCCGCTGGGCAAGGACTCGCCCTCCAAACAGCTGTGTAT	5400
QY	5401	GCCAAAGACATATCCCACTCTCAAGAATTTGGTGTGAGAGATTAATCTAGACATATAGGAAG	5460
Db	5401	GCCAAAGACATATCCCACTCTCAAGAATTTGGTGTGAGAGATTAATCTAGACATATAGGAAG	5460
QY	5461	ATGCGACCCCATACAGCAACCAAGCATATACCTGTGCTGAGCAGTCCCGATGAC	5520
Db	5461	ATGCGACCCCATACAGCAACCAAGCATATACCTGTGCTGAGCAGTCCCGATGAC	5520
QY	5521	ATGATATAGTTTCAACACCATGAGAGCACTCTCAAGAGATCTTTCTCTATGTGGGCAATAC	5580
Db	5521	ATGATATAGTTTCAACACCATGAGAGCACTCTCAAGAGATCTTTCTCTATGTGGGCAATAC	5580
QY	5581	AGCAGAGAGATCTCTTGAGCCTCTGACCAACATGACCCAGTGTGGGAAGCAAACTGGCC	5640
Db	5581	AGCAGAGAGATCTCTTGAGCCTCTGACCAACATGACCCAGTGTGGGAAGCAAACTGGCC	5640
QY	5641	TACAACTATAGAACATGATTAACCTTCATAGCTTTACAGCTGAATATAA	5691
Db	5641	TACAACTATAGAACATGATTAACCTTCATAGCTTTACAGCTGAATATAA	5691
RESULT 2			
LOCUS	AX492982	6367 bp	DNA
DEFINITION	Sequence 19 from Patent WO02059212.	Linear	PAT 26-SEP-2002
ACCESSION	AX492982		
VERSION	AX492982.1	GI:23338637	
KEYWORDS			
SOURCE			
ORGANISM	Homo sapiens (human)		
	Homo sapiens		
	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;		
	Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo.		
REFERENCE			
AUTHORS	1. Kalliock D.A., Lee S., Xu Y., Yao M.G., Yue H., Bandman O.B.,		
	Buford N., Gandel A.R., Grall R.C., Gal P.G., Lu D.A., Lu Y.,		
	Tang T.Y., Duggan B.M., Gietzen K.O., Hillman J.V., Honchell C.D.,		
	Ramkumar J., Wallis N.K. and Warren B.A.		
	Cell adhesion proteins		
	Patent: WO 02059312-A 19 01-AUG-2002;		
JOURNAL	INCYTE GENOMICS INC (US)		
FEATURES			
source	1..6367		
	/organism="Homo sapiens"		
	/mol_type="unassigned DNA"		
	/db_xref="taxon:9606"		
	/note="Incyte ID No: 7156379CE1"		
ORIGIN			
Query Match	99.8%; Score 5680.6; DB 6; Length 6367;		
Best Local Similarity	99.9%; Pred. No. 0;		
Matches 5683; Conservative	0; Mismatches 4; Indels 0; Gaps 0;		
QY	1	ATGAAAGCATGCCCTGGAACTGGAAGCTGCTTCTCTCCAACTCTCATGTGGGATG	60
Db	566	ATGAAAGCATGCCCTGGAACTGGAAGCTGCTTCTCTCCAACTCTCATGTGGGATG	625
QY	61	GAGCTCTCACTTGTGTCAACCGGACGACCCCGCTGTCCAGAAAGACGAGTCAATT	120
Db	626	GAGCTCTCACTTGTGTCAACCGGACGACCCCGCTGTCCAGAAAGACGAGTCAATT	685
QY	121	GTCACTTCCGAGAGAGCCCGCGAGAGGTTTCAATCACTGTGTGTGATATGAGACA	180
Db	686	GTCACTTCCGAGAGAGCCCGCGAGAGGTTTCAATCACTGTGTGTGATATGAGACA	745

181 GGACACATTTACTTGGGGGCGCTGATCGGATTTTACAGCTTCCAGCGACCTGAGTTC 240
746 GGACACATTTACTTGGGGGCGCTGATCGGATTTTACAGCTTCCAGCGACCTGAGTTC 805
241 TTGGTGAAGCATGAGACAGGGCCGAGAGACAAACCCCAAGTGTATCCCAACCCGATC 300
806 TTGGTGAAGCATGAGACAGGGCCGAGAGACAAACCCCAAGTGTATCCCAACCCGATC 865
301 GTCCAGACCTTGCATGAGCCCTTCAACCAACCAATGTCAACAGATGCTCTCATTA 360
866 GTCCAGACCTTGCATGAGCCCTTCAACCAACCAATGTCAACAGATGCTCTCATTA 925
361 GACTTCAAGAGGAAAGAGGCTGATGCTGATGAGGAGCTGTACCAAGGATCTGCAAGCTG 420
926 GACTTCAAGAGGAAAGAGGCTGATGCTGATGAGGAGCTGTACCAAGGATCTGCAAGCTG 985
421 CTGAGGCTGAGAGACCTCTTCAAGCTGAGGAGGCTTATCATTAAGAGAGCACTATCTG 480
986 CTGAGGCTGAGAGACCTCTTCAAGCTGAGGAGGCTTATCATTAAGAGAGCACTATCTG 1045
481 TCAGGTGTCAAGAGAGGAGGCTGATGCTGATGAGGAGCTGTACCAAGGATCTGCAAGCTG 540
1046 TCAGGTGTCAAGAGAGGAGGCTGATGCTGATGAGGAGCTGTACCAAGGATCTGCAAGCTG 1105
541 GACAGAGCTGTTCATTTGCCACGCGAGTGGATGGAAAGCCCGAGTATTTTCCACCATCTCC 600
1106 GACAGAGCTGTTCATTTGCCACGCGAGTGGATGGAAAGCCCGAGTATTTTCCACCATCTCC 1165
601 AGCCGGAACCTGACCAAGAAATCTGAGGCGGATGGCAATGTTCCGGTACGGTTCGATGAT 660
1166 AGCCGGAACCTGACCAAGAAATCTGAGGCGGATGGCAATGTTCCGGTACGGTTCGATGAT 1225
661 GAGTTCGTCGCTCGATGATTAAGATCCCTTCGAGCACTTCACCATCATCTCGATCTTT 720
1226 GAGTTCGTCGCTCGATGATTAAGATCCCTTCGAGCACTTCACCATCATCTCGATCTTT 1285
721 GATATCTATCTATGCTATGCTTTTACAGAGTGGCACTTTGCTACTTTTGAACCTTCGAA 780
1286 GATATCTATCTATGCTATGCTTTTACAGAGTGGCACTTTGCTACTTTTGAACCTTCGAA 1345
781 CCTGAGATGCTGTCTCCACAGGCTCCACACCAAGAGAGAGTGTATATACATCCAGCTC 840
1346 CCTGAGATGCTGTCTCCACAGGCTCCACACCAAGAGAGAGTGTATATACATCCAGCTC 1405
841 GTGAGGCTTTGCAAGAGAGAGACAGGCTTCACTCTATGTAAGAGTGCCTATTTGGCTGT 900
1406 GTGAGGCTTTGCAAGAGAGAGACAGGCTTCACTCTATGTAAGAGTGCCTATTTGGCTGT 1465
901 GAGGCGAAGGGGTGAGATGACCGGCTGCTGAGAGGCTGCTACCTGTCMAAAGCGGAGGC 960
1466 GAGGCGAAGGGGTGAGATGACCGGCTGCTGAGAGGCTGCTACCTGTCMAAAGCGGAGGC 1525
961 GTGCTTGGCAGAGACCTTGAATCCATCAAGATGATGATCTGCTCTTCAACCGTCTTCTCC 1020
1526 GTGCTTGGCAGAGACCTTGAATCCATCAAGATGATGATCTGCTCTTCAACCGTCTTCTCC 1585
1021 AAGGCGCAGAGCGGAGAAATGAAATCCCTGATGATGAGGCGCTGATCTTCAATCTG 1080
1586 AAGGCGCAGAGCGGAGAAATGAAATCCCTGATGATGAGGCGCTGATCTTCAATCTG 1645
1081 AAGCAGATTAATGACCGCATTAAGAGAGGCTGAGTCTTGTATCCGAGGAGAGAGCAGC 1140
1646 AAGCAGATTAATGACCGCATTAAGAGAGGCTGAGTCTTGTATCCGAGGAGAGAGCAGC 1705
1141 CTGAGACCTTGGCTGCTCAAGGTGAAGACATCCCTGAGAGGCTCTTAAACATTT 1200
1706 CTGAGACCTTGGCTGCTCAAGGTGAAGACATCCCTGAGAGGCTCTTAAACATTT 1765
1201 GACGATTAATCTTGTGAGCTGAGACATGATGCTTCCCTGAGAGTGTCCGACATGGTGGCT 1260
1766 GACGATTAATCTTGTGAGCTGAGACATGATGCTTCCCTGAGAGTGTCCGACATGGTGGCT 1825

1261 GGAATTTCCGCTCTTCCAGGAGAGACAGGACCGCATGACGCTGTGATGATATGTCTAC 1320
1826 GGAATTTCCGCTCTTCCAGGAGAGACAGGACCGCATGACGCTGTGATGATATGTCTAC 1885
1321 AAGAACCACTCTTGGCTTTGTGGGACCAAAAGTGGCAAGCTGAAAGATCCGGGTG 1380
1886 AAGAACCACTCTTGGCTTTGTGGGACCAAAAGTGGCAAGCTGAAAGATCCGGGTG 1945
1381 GATGAGACCAAGGAGCAAGCCCTCCAGATGAGACGATGAGTGTGTGACCCCGCCA 1440
1946 GATGAGACCAAGGAGCAAGCCCTCCAGATGAGACGATGAGTGTGTGACCCCGCCA 2005
1441 GTCTCTCCGAGATATGAGCTTCTCCAGAGACCAAGCACTCTATCATATGTCAAGAG 1500
2006 GTCTCTCCGAGATATGAGCTTCTCCAGAGACCAAGCACTCTATCATATGTCAAGAG 2065
1501 GAGCTCACAGAGTCCCTGTGAGTCTGTGAGTATCAGAGCTGCGAGTGCCTT 1560
2066 GAGCTCACAGAGTCCCTGTGAGTCTGTGAGTATCAGAGCTGCGAGTGCCTT 2125
1561 GAGCTCACAGAGTCCCTGTGAGTCTGTGAGTATCAGAGCTGCGAGTGCCTT 1620
2126 GAGCTCACAGAGTCCCTGTGAGTCTGTGAGTATCAGAGCTGCGAGTGCCTT 2185
1621 GAGCTCACAGAGTCCCTGTGAGTCTGTGAGTATCAGAGCTGCGAGTGCCTT 1680
2186 GAGCTCACAGAGTCCCTGTGAGTCTGTGAGTATCAGAGCTGCGAGTGCCTT 2245
1681 GAGCTCACAGAGTCCCTGTGAGTCTGTGAGTATCAGAGCTGCGAGTGCCTT 1740
2246 GAGCTCACAGAGTCCCTGTGAGTCTGTGAGTATCAGAGCTGCGAGTGCCTT 2305
1741 GAGCTCACAGAGTCCCTGTGAGTCTGTGAGTATCAGAGCTGCGAGTGCCTT 1800
2306 GAGCTCACAGAGTCCCTGTGAGTCTGTGAGTATCAGAGCTGCGAGTGCCTT 2365
1801 GAGCTCACAGAGTCCCTGTGAGTCTGTGAGTATCAGAGCTGCGAGTGCCTT 1860
2366 GAGCTCACAGAGTCCCTGTGAGTCTGTGAGTATCAGAGCTGCGAGTGCCTT 2425
1861 GAGCTCACAGAGTCCCTGTGAGTCTGTGAGTATCAGAGCTGCGAGTGCCTT 1920
2426 GAGCTCACAGAGTCCCTGTGAGTCTGTGAGTATCAGAGCTGCGAGTGCCTT 2485
1921 GAGCTCACAGAGTCCCTGTGAGTCTGTGAGTATCAGAGCTGCGAGTGCCTT 1980
2486 GAGCTCACAGAGTCCCTGTGAGTCTGTGAGTATCAGAGCTGCGAGTGCCTT 2545
1981 GAGCTCACAGAGTCCCTGTGAGTCTGTGAGTATCAGAGCTGCGAGTGCCTT 2040
2546 GAGCTCACAGAGTCCCTGTGAGTCTGTGAGTATCAGAGCTGCGAGTGCCTT 2605
2041 GAGCTCACAGAGTCCCTGTGAGTCTGTGAGTATCAGAGCTGCGAGTGCCTT 2100
2606 GAGCTCACAGAGTCCCTGTGAGTCTGTGAGTATCAGAGCTGCGAGTGCCTT 2665
2666 GAGCTCACAGAGTCCCTGTGAGTCTGTGAGTATCAGAGCTGCGAGTGCCTT 2725
2101 GAGCTCACAGAGTCCCTGTGAGTCTGTGAGTATCAGAGCTGCGAGTGCCTT 2160
2726 GAGCTCACAGAGTCCCTGTGAGTCTGTGAGTATCAGAGCTGCGAGTGCCTT 2785
2221 GAGCTCACAGAGTCCCTGTGAGTCTGTGAGTATCAGAGCTGCGAGTGCCTT 2845
2786 GAGCTCACAGAGTCCCTGTGAGTCTGTGAGTATCAGAGCTGCGAGTGCCTT 2905
2281 GAGCTCACAGAGTCCCTGTGAGTCTGTGAGTATCAGAGCTGCGAGTGCCTT 2960
2846 GAGCTCACAGAGTCCCTGTGAGTCTGTGAGTATCAGAGCTGCGAGTGCCTT 3020
2341 GAGCTCACAGAGTCCCTGTGAGTCTGTGAGTATCAGAGCTGCGAGTGCCTT 3080

Dp	2906	TTGACAGTGTGTGGAGTGGGCACTTCAACATTGACAAACCAGCTCGAATTAAGTTAC	2965
Qy	2401	CTCTACAAAGTGTGGAGCCATGCGGTGAGAGTCTGGCGGGTGTGCTCAAGGGTGAACCCAGAC	2460
Dp	2966	CTTCAACAAGTGTGGAGCCATGCGGTGAGAGTCTGGCGGGTGTGCTCAAGGGTGAACCCAGAC	3025
Qy	2461	TTGCGCATGTGGCTGGTGTCCAGGGGCCAGGGCCAGTGCACCTGTGGCCAGCACTGCGCTGCC	2520
Dp	3026	TTGCGCATGTGGCTGGTGTCCAGGGGCCAGGGCCAGTGCACCTGTGGCCAGCACTGCGCTGCC	3085
Qy	2521	CAGAGAGCCAGTGGTGTGGAGCTGTCTGTGTGCCAAAGTGTGCACAAACCCCGCATC	2580
Dp	3086	CAGAGAGCCAGTGGTGTGGAGCTGTCTGTGTGCCAAAGTGTGCACAAACCCCGCATC	3145
Qy	2581	ACAGAGATTAATCCGCGTGTGACAGAGCCCGGGGAAGGGGGACACAGTCACTATCCAGGG	2640
Dp	3146	ACAGAGATTAATCCGCGTGTGACAGAGCCCGGGGAAGGGGGACACAGTCACTATCCAGGG	3205
Qy	2641	GAGAACTTGGGCTTGGAAATTTCCGACATGCGCTCCATGTCAAGTTGCTGGCGTGGAG	2700
Dp	3206	GAGAACTTGGGCTTGGAAATTTCCGACATGCGCTCCATGTCAAGTTGCTGGCGTGGAG	3265
Qy	2701	TGCAGCCCTTAAAGTGAATGTTACATCCCTTGACAAACAGATCGTGTGTGAGATGGGGAG	2760
Dp	3266	TGCAGCCCTTAAAGTGAATGTTACATCCCTTGACAAACAGATCGTGTGTGAGATGGGGAG	3325
Qy	2761	GCCAAAGCCAGCCAGCATGACAGCTTGTGGAGATCTGCGTGGTGTGTGTCCGCTCGAA	2820
Dp	3326	GCCAAAGCCAGCCAGCATGACAGCTTGTGGAGATCTGCGTGGTGTGTGTCCGCTCGAA	3385
Qy	2821	TTCAATGGCCGAGTCTCTCAGCTCTATTAATCTTCAACATGCACTCTCTCAGATCTTGAG	2880
Dp	3386	TTCAATGGCCGAGTCTCTCAGCTCTATTAATCTTCAACATGCACTCTCTCAGATCTTGAG	3445
Qy	2881	CCCAAGCCGGGGGCCCATATGTCCGGAGGGACCCAGTATACATCAAGGCACTAAGCTTGAT	2940
Dp	3446	CCCAAGCCGGGGGCCCATATGTCCGGAGGGACCCAGTATACATCAAGGCACTAAGCTTGAT	3505
Qy	2941	GCCGGAACCAAGTGTGTGATGTTTGGAAAGCAGCCCTGTCTCTTCCAGCGCATCT	3000
Dp	3506	GCCGGAACCAAGTGTGTGATGTTTGGAAAGCAGCCCTGTCTCTTCCAGCGCATCT	3565
Qy	3001	CCATCTTAACATTTGTCTGCACACCATCTCTCAGATGAGTGTCTAGATTAAGATGTGTGCG	3060
Dp	3566	CCATCTCTCATTTGTCTGCACACCATCTCTCAGATGAGTGTCTAGATTAAGATGTGTGCG	3625
Qy	3061	GTGCAAGTGTGACAGGGCCCAAGTCCACACAGGACCTGTCTTCAAGTATGTGAAGACCCC	3120
Dp	3626	GTGCAAGTGTGACAGGGCCCAAGTCCACACAGGACCTGTCTTCAAGTATGTGAAGACCCC	3685
Qy	3121	ACCATCTGTGCGGATTGAGCCAGATGTGAGCATTTGTAGTGAACACACCAATGGCGCTA	3180
Dp	3686	ACCATCTGTGCGGATTGAGCCAGATGTGAGCATTTGTAGTGAACACACCAATGGCGCTA	3745
Qy	3181	TGGGGGACCCACCTTGACCTTCAACAGAAACCCCGAGATCCGCGCAACATGTGAGGGAG	3240
Dp	3746	TGGGGGACCCACCTTGACCTTCAACAGAAACCCCGAGATCCGCGCAACATGTGAGGGAG	3805
Qy	3241	GAGCACAATCAATATCTGTGAGGTTCTGAACGCTACTGAGATGACTTGTGAGCGCGCCG	3300
Dp	3806	GAGCACAATCAATATCTGTGAGGTTCTGAACGCTACTGAGATGACTTGTGAGCGCGCCG	3865
Qy	3301	CTGCGCTCTGGGTCTTGACCAACAGTCAAGCTGACCGAGAGGCGCCGAGAGATTGGCTTC	3360
Dp	3866	CTGCGCTCTGGGTCTTGACCAACAGTCAAGCTGACCGAGAGGCGCCGAGAGATTGGCTTC	3925
Qy	3361	ATCTGTGACAAAGTGTGAGTGTGCTGTCACTCTCAACAGACCACTTCACTCACTATCTCC	3420
Dp	3926	ATCTGTGACAAAGTGTGAGTGTGCTGTCACTCTCAACAGACCACTTCACTCACTATCTCC	3985
Qy	3421	AATCCGATGTTGAGGCTTTGGTCTCTCAGGATCTTGAAGCTCAAGCTTGACGCGCC	3480

[illegible]

QY	4561	AACGTGACACCAATCACTCAAGGTCAAGAGAAAGTTCTGGATGCGATCTTCAAGAAATG	4620
Db	5126	AACTGTACACCACTCACTCAAGTCAAGAGAAAGTTCTGGATGCGATCTTCAAGAAATG	5185
QY	4621	CTTTGCTCCACCGGCCCCAAGCTGCAGATATGGATCTGGAGTGGCGACAAGAAATGG	4680
QY	5186	CTTTGCTCCACCGGCCCCAAGCTGCAGATATGGATCTGGAGTGGCGACAAGAAATGG	5245
Db	4681	GCAAGATGATCTTGCAGAGATGAAGACATCAACCAACAAATTGAATGATTTGAAGCGA	4740
QY	5246	GCAAGATGATCTTGCAGAGATGAAGACATCAACCAACAAATTGAATGATTTGAAGCGA	5305
QY	4741	CTGAACAACACTGGCCCACTACCAAGTCCAGATGTTCCGTGGTGGCATTAATGTCCAG	4800
Db	5306	CTGAACAACACTGGCCCACTACCAAGTCCAGATGTTCCGTGGTGGCATTAATGTCCAG	5365
QY	4801	CAGGTGACAGCTCTTAACGACAGTGAACAACTCCACCGTCTCAAGGACCTCAGACAATAA	4860
Db	5366	CAGGTGACAGCTCTTAACGACAGTGAACAACTCCACCGTCTCAGGACCTCAGACAATAA	5425
QY	4861	TATGAAAAATGATCCGGTACACCGGCGACGCCCGACACGCTCCGCTCAGCGACACTATG	4920
Db	5426	TATGAAAAATGATCCGGTACACCGGCGACGCCCGACACGCTCCGCTCAGCGACACTATG	5485
QY	4921	ATCACTCTTGACTCTGGAGAGTGAAGTCAAGATGTGGACCTAGTGAAGAACCAAGAC	4980
Db	5486	ATCACTCTTGACTCTGGAGAGTGAAGTCAAGATGTGGACCTAGTGAAGAACCAAGAC	5545
QY	4981	GGAGACCAAGAGAGGGGGACCGGGGGAGCAAGATGTTGTGAATCTTACCTGACCCGA	5040
Db	5546	GGAGACCAAGAGAGGGGGACCGGGGGAGCAAGATGTTGTGAATCTTACCTGACCCGA	5605
QY	5041	CTCTGGGCACTAAGGGGACACTCAGAAAGTTGTGATGACTCTTTGAGAACCATCTTC	5100
Db	5606	CTCTGGGCACTAAGGGGACACTCAGAAAGTTGTGATGACTCTTTGAGAACCATCTTC	5665
QY	5101	AGCAGCGCACCCGTGCTCTGCTCCCTGCGCATCATGTAATGTTGAATCTTCTG	5160
Db	5666	AGCAGCGCACCCGTGCTCTGCTCCCTGCGCATCATGTAATGTTGAATCTTCTG	5725
QY	5161	GATGAGCAGGCTGATTAACATGTCATACCGGACGTCGCCCATACTCTGAGAAAGC	5220
Db	5726	GATGAGCAGGCTGATTAACATGTCATACCGGACGTCGCCCATACTCTGAGAAAGC	5785
QY	5221	AATTGCTGCCCCGTAGGTTTGGGTCAATGATCAAGAACCCGCACTTGTGTGGAC	5280
Db	5786	AATTGCTGCCCCGTAGGTTTGGGTCAATGATCAAGAACCCGCACTTGTGTGGAC	5845
QY	5281	ATCCATAAGAACAGATCAACAGACGCTCGCTCTGTGTGGCTCAGACCTTCATGGAC	5340
Db	5846	ATCCATAAGAACAGATCAACAGACGCTCGCTCTGTGTGGCTCAGACCTTCATGGAC	5905
QY	5341	TCTTGCTCAAGTCAAGACACCGGTGGGCAAGACTCGCCCTCCAAACACTGTGAT	5400
Db	5906	TCTTGCTCAAGTCAAGACACCGGTGGGCAAGACTCGCCCTCCAAACACTGTGAT	5965
QY	5401	GCCAAAGACATCCCGAGCTACAGAAATTGGGTGAGAGTATTACTCAGATAGGGAAG	5460
Db	5966	GCCAAAGACATCCCGAGCTACAGAAATTGGGTGAGAGTATTACTCAGATAGGGAAG	6025
QY	5461	ATGCCAGGCATAGCGACCAAGACATGAACGATCTCTGGCTGAGCACTCCCGATGGAC	5520
Db	6026	ATGCCAGGCATAGCGACCAAGACATGAACGATCTCTGGCTGAGCACTCCCGATGGAC	6085
QY	5521	ATGAATGAGTTCAACACCATGAGTCACTCTCAGAGATCTTCTCCTATGTGGGCAAAATAC	5580
Db	6086	ATGAATGAGTTCAACACCATGAGTCACTCTCAGAGATCTTCTCCTATGTGGGCAAAATAC	6145
QY	5581	AGCGAGAGATCTCTTGAACCTTGAACCAAGATGACAGTGTGGGAACCAAGAACTGGCC	5640
Db	6146	AGCGAGAGATCTCTTGAACCTTGAACCAAGATGACAGTGTGGGAACCAAGAACTGGCC	6205

[illegible]

ORIGIN MSALSELFVYVKYSEELIPLHDDQCKOKLAKLEQVITLMSLDS"

Query Match 82.8%; Score 4714.2; DB 10; Length 7022;
Best Local Similarity 89.4%; Pred. No. 0;
Matches 5086; Conservative 0; Mismatches 598; Indels 3; Gaps 1;

QY 1 ATGAAAGCATGCGCTGAGTCACTGACCTGCTCTCCACCACTCCATGATGAGGATG 60
DB ATGAAAGCATGCGCTGAGTCACTGACCTGCTCTCCACCACTCCATGATGAGGATG 1093
QY 974 ATGAAAGCATGCGCTGAGTCACTGACCTGCTCTCCACCACTCCATGATGAGGATG 1093
DB 61 GGCCTCTCCACCTGCTCTCCACCACTCCATGATGAGGATG 120
QY 1034 GGCCTCTCCACCTGCTCTCCACCACTCCATGATGAGGATG 1090
DB 121 GTCATATTCGAG 180
QY 1091 GTCATATTCGAG 1150
DB 181 GGCATATTCGAG 240
QY 1151 GGCATATTCGAG 1210
DB 241 TTGAGTACGATGAG 300
QY 1211 TTGAGTACGATGAG 1270
DB 301 GTCCAGACCTGATGAG 360
QY 1271 GTCCAGACCTGATGAG 1330
DB 361 GATCTCAAG 420
QY 1331 GATCTCAAG 1390
DB 421 CTGAGAGCTGAG 480
QY 1391 CTGAGAGCTGAG 1450
DB 481 TCGAGTGTCAAG 540
QY 1451 TCGAGTGTCAAG 1510
DB 541 GATCAAGCTGATGAG 600
QY 1511 GATCAAGCTGATGAG 1570
DB 601 AGCCGAGAACTGACCAAG 660
QY 1571 AGCCGAGAACTGACCAAG 1630
DB 661 GAGTTCGTGAGCTGATGATTAAGATCCCTTCGAGACCTTCACCATATCCGAGACTT 720
QY 1631 GAGTTCGTGAGCTGATGATTAAGATCCCTTCGAGACCTTCACCATATCCGAGACTT 1690
DB 721 GATATCTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 780
QY 1691 GATATCTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1750
DB 781 CCTGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 840
QY 1751 CCTGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1810
DB 841 GTGAGGCTTTGCAAG 900
QY 1811 GTGAGGCTTTGCAAG 1870
DB 901 GAGCGGAGTGGGAGTGAAGTACCGCTGCTGAGAGAGAGAGAGAGAGAGAGAGAGAG 960
QY 1871 GAGCGGAGTGGGAGTGAAGTACCGCTGCTGAGAGAGAGAGAGAGAGAGAGAGAGAG 1930
DB 961 GTGCTTGGCAGAGACCTTGGAGTGCATCAGATGATGATGATGATGATGATGATGAT 1020

DB 1931 GTGCTTGGCAGAGACCTTGGAGTGCATCAGATGATGATGATGATGATGATGATGAT 1990
QY 1021 AAGGCGCAG 1080
DB 1991 AAGGCGCAG 2050
QY 1081 AAGCAGATTAATGACCGAGATTAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1140
DB 2051 AAGCAGATTAATGACCGAGATTAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2110
QY 1141 CTGAGACTGAGCTGAGTCAAGTGAAGAGATCCCTTCAGAGAGAGAGAGAGAGAGAG 1200
DB 2111 CTGAGAGCTGAGCTGAGTCAAGTGAAGAGATCCCTTCAGAGAGAGAGAGAGAGAG 2170
QY 1201 GACGATTAATCTGAGCTGAGTCAAGTGAAGAGATCCCTTCAGAGAGAGAGAGAGAG 1260
DB 2171 GATGACAACTTGTGAGCTGAGTCAAGTGAAGAGATCCCTTCAGAGAGAGAGAGAGAG 2230
QY 1261 GGAATTCCTGCTTTCAAG 1320
DB 2231 GGCATTCAGATCTTCAAG 2290
QY 1321 AAGAACAATCTCTGAGCTTGTGAGGACCAAAAGTGGCAAGCTGAAGAGAGATCCGAGT 1380
DB 2291 AAGAACAATCTCTGAGCTTGTGAGGACCAAAAGTGGCAAGCTGAAGAGAGATCCGAGT 2350
QY 1381 GATGACCAAG 1440
DB 2351 GATGAGGAG 2410
QY 1441 GTCCCTCGGAGATGAGCTTGTGAGGACCAAAAGTGGCAAGCTGAAGAGAGATCCGAGT 1500
DB 2411 GTCCCTCGGAGATGAGCTTGTGAGGACCAAAAGTGGCAAGCTGAAGAGAGATCCGAGT 2470
QY 1501 CAGCTCAAG 1560
DB 2471 CAGCTCAAG 2530
QY 1561 GAGTCAAG 1620
DB 2531 GAGTCAAG 2590
QY 1621 CGGTGTGAG 1680
DB 2591 CGGTGTGAG 2650
QY 1681 CTGACGATCCATCCCAAG 1740
DB 2651 CTGACGATCCATCCCAAG 2710
QY 1741 ACATACATGATCCCGAGAGTGTGAGTGGAGTCAAGTCAAGTTCAGAGAGAGAGAG 1800
DB 2711 ACATACATGATCCCGAGAGTGTGAGTGGAGTCAAGTCAAGTTCAGAGAGAGAGAG 2770
QY 1801 ATGAGAGAGTGTGAGTGGAGTCAAGTTCAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1860
DB 2771 ATGAGAGAGTGTGAGTGGAGTCAAGTTCAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2830
QY 1861 CCCGAGATCAACAG 1920
DB 2831 CCCGAGATCAACAG 2890
QY 1921 GAGAGCGGATGAGTGTGAGTGGAGTCAAGTTCAGAGAGAGAGAGAGAGAGAGAGAGAG 1980
DB 2891 GAGAGCGGATGAGTGTGAGTGGAGTCAAGTTCAGAGAGAGAGAGAGAGAGAGAGAGAG 2950
QY 1981 TCGTGCCTGCTGCTGAGAGAGTCCATCCCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2040
DB 2951 TCGTGCCTGCTGCTGAGAGAGTCCATCCCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 3010
QY 2041 TCGACCAATGAG 2100
DB 3011 TCGACCAATGAG 3070

QY 2101 TGCCCCAGCTGTCGAGTGAACAAGATCTGGTGGCCGTGAGGTGATCAAGCTATC 2160
DB 3071 TGCCCCAGCTGTCGAGTGAACAAGATCTGGTGGCCGTGAGGTGATCAAGCTATC 3130
QY 2161 AGCTGAAGGCGAAGAACCTTCCCGAGCCCAAGTCTGGGCAAGCTGGCTACGAATGATC 2220
DB 3131 ACTCTGAAGGCTAAGAACCTTCCCGAGCCCAAGTCTGGGCAAGAGGCTACGAATGATC 3190
QY 2221 CTCAACATTCAGGGCAGGAGCAGAGAGTGGCCGCTGGGCTTCAACAGCTCCAGGTA 2280
DB 3191 TTGAACATTCAGGGCAGGAGCAGAGAGTGGCTGGGCTTCAACAGCTCCAGGTA 3250
QY 2281 CAGTGCAGAACACCTCTTATTCCTATGAAGGAGTGAATCAACACCTGCTGGAG 2340
DB 3251 CAGTGCAGAACACCTCTTATTCCTATGAAGGAGTGAATCAACACCTGCTGGAG 3310
QY 2341 TTGAACATTCAGGGCAGGAGCAGAGAGTGGCCGCTGGGCTTCAACAGCTCCAGGTA 2400
DB 3311 TTGAACATTCAGGGCAGGAGCAGAGAGTGGCCGCTGGGCTTCAACAGCTCCAGGTA 3370
QY 2401 CTCTACAGTGTGAGGAGCAGTGCAGAGAGTGGCCGCTGGGCTTCAACAGCTCCAGGTA 2460
DB 3371 CTCTACAGTGTGAGGAGCAGTGCAGAGAGTGGCCGCTGGGCTTCAACAGCTCCAGGTA 3430
QY 2461 TTGGCAGTGTGAGGAGCAGTGCAGAGAGTGGCCGCTGGGCTTCAACAGCTCCAGGTA 2520
DB 3431 TTGGCAGTGTGAGGAGCAGTGCAGAGAGTGGCCGCTGGGCTTCAACAGCTCCAGGTA 3490
QY 2521 CAGAGAGAGCAGTGTGAGGAGCAGTGCAGAGAGTGGCCGCTGGGCTTCAACAGCTCCAGGTA 2580
DB 3491 CAGAGAGAGCAGTGTGAGGAGCAGTGCAGAGAGTGGCCGCTGGGCTTCAACAGCTCCAGGTA 3550
QY 2581 ACAAGATATATCCCGGTGACAGGCCCCCGGGAAGGGGACCAAGTCACTATTCGAGG 2640
DB 3551 ACAAGATATATCCCGGTGACAGGCCCCCGGGAAGGGGACCAAGTCACTATTCGAGG 3610
QY 2641 GAGAAGCCTGGGCTGGAAATTTCCGAGCATGCTCCCTCCATGTCAGATGGCTGGAG 2700
DB 3611 GAGAAGCCTGGGCTGGAAATTTCCGAGCATGCTCCCTCCATGTCAGATGGCTGGAG 3670
QY 2701 TGCAAGCCTTATGAGATGTTACATCCCTGCAAGACAGATGCTGTGTGAGATGGGAG 2760
DB 3671 TGCAAGCCTTATGAGATGTTACATCCCTGCAAGACAGATGCTGTGTGAGATGGGAG 3730
QY 2761 GCGAAGCCCAAGCAGATGAGAGCTTGGTGAAGATGCTGTGTGAGATGGGAG 2820
DB 3731 GCGAAGCCCAAGCAGATGAGAGCTTGGTGAAGATGCTGTGTGAGATGGGAG 3790
QY 2821 TTCAATGAGCCTGCTCAAGCTCTATTAATTCATGACATGACTCTCTCAATCTGAG 2880
DB 3791 TTCAATGAGCCTGCTCAAGCTCTATTAATTCATGACATGACTCTCTCAATCTGAG 3850
QY 2881 CCCAGCCGGGGGCGCATGCTCCGAGAGGACCCCAATGACATCAAGGACCAACCTGAAT 2940
DB 3851 CCCAGCCGGGGGCGCATGCTCCGAGAGGACCCCAATGACATCAAGGACCAACCTGAAT 3910
QY 2941 GCGGAGAACAGATGAGTGTGATGTTGGAAGCAGCCCTGCTCTTCAAGGAGTCT 3000
DB 3911 GCGGAGAACAGATGAGTGTGATGTTGGAAGCAGCCCTGCTCTTCAAGGAGTCT 3970
QY 3001 CCATCTTCAATTTGTCGACACCAATCTCAAGTGAAGTGTGAGATGAGATGAGTGTG 3060
DB 3971 CCATCTTCAATTTGTCGACACCAATCTCAAGTGAAGTGTGAGATGAGATGAGTGTG 4030
QY 3061 GTGAGGTGACAGGGGCCAAGATCCACAGGACCTGCTCTTCAAGTGTGAGAGCC 3120
DB 4031 GTGAGGTGACAGGGGCCAAGATCCACAGGACCTGCTCTTCAAGTGTGAGAGCC 4090
QY 3121 ACCATGCTGGGATTTGAGCCAGATGAGCATTTGTCACTGAGAACACACCATGCTG 3180
DB 4091 ACCATGCTGGGATTTGAGCCAGATGAGCATTTGTCACTGAGAACACACCATGCTG 4150

QY 3181 TGGGGAGCCCACTTGACCTCATACAGAACCCCGAGATCCGTGCTCAAGCATGAGGAG 3240
DB 4151 TGGGGAGCCCACTTGACCTCATACAGAACCCCGAGATCCGTGCTCAAGCATGAGGAG 4210
QY 3241 GAGCATCAATATGAGAGGTTCTGAAGCTCATGAGATGAGACCTGTCAGGCGCCG 3300
DB 4211 GAGCATCAATATGAGAGGTTCTGAAGCTCATGAGATGAGACCTGTCAGGCGCCG 4270
QY 3301 CTGCTCTGGGCTCTGACCAAGCAGTCAAGCTGACCGAGAGGCGGAGAGTTGGCTTC 3360
DB 4271 CTGCTCTGGGCTCTGACCAAGCAGTCAAGCTGACCGAGAGGCGGAGAGTTGGCTTC 4330
QY 3361 ATCTGAGACAGTGCATGCTCTGATCTCTCAACAGACCAATTCATCTACATCC 3420
DB 4331 ATCTGAGACAGTGCATGCTCTGATCTCTCAACAGACCAATTCATCTACATCC 4390
QY 3421 AACCGGATTTGAGGCTTGGTCCCTCAGGAATCCTGAGCTCAAGCTGAGCGCC 3480
DB 4391 AACCGGATTTGAGGCTTGGTCCCTCAGGAATCCTGAGCTCAAGCTGAGCGCC 4450
QY 3481 ATCATCTTAAGGCAAGACCTGATCCCGCTGTGGCTGGGAGCAAGTGAAGCTGAC 3540
DB 4451 ATCATCTTAAGGCAAGACCTGATCCCGCTGTGGCTGGGAGCAAGTGAAGCTGAC 4510
QY 3541 TACATGCTGTGTTGGGAGAGAGCGTGAACCGTGAACCGTGTGATGTCCAGTGTCTC 3600
DB 4511 TACATGCTGTGTTGGGAGAGAGCGTGAACCGTGAACCGTGTGATGTCCAGTGTCTC 4570
QY 3601 TGCAGTCCCACTCATGCGCAGGACCAAGGATGAGCAAGGATGAGTGGAGTGGAG 3660
DB 4571 TGCAGTCCCACTCATGCGCAGGACCAAGGATGAGCAAGGATGAGTGGAGTGGAG 4630
QY 3661 TACTCCCGGGAGATGATGATCAATGCCCCGGAACACCGCTGAGCTGCGGCAATGCTC 3720
DB 4631 TACTCCCGGGAGATGATGATCAATGCCCCGGAACACCGCTGAGCTGCGGCAATGCTC 4690
QY 3721 AGCATGAGTGGCTGGGAGCTCTCTCATATTTTCATGATGAGCCGTGATGCTAT 3780
DB 4691 AGCATGAGTGGCTGGGAGCTCTCTCATATTTTCATGATGAGCCGTGATGCTAT 4750
QY 3781 AAAAGCAATCCCGGAAAGTAACTCAAGCTGAAGCGGCTGCAATGAGTGAAGCAAC 3840
DB 4751 AAAAGCAATCCCGGAAAGTAACTCAAGCTGAAGCGGCTGCAATGAGTGAAGCAAC 4810
QY 3841 CTGAGATCCCGTGGGCTCTGAGTGAAGAGCCTTTCAGGCTGAGAGCAATC 3900
DB 4811 CTGAGATCCCGTGGGCTCTGAGTGAAGAGCCTTTCAGGCTGAGAGCAATC 4870
QY 3901 CATGAGCTGACAGTGAAGCTGAGTGAAGCGGAGATTCGTTCTGAGATGAGATGAG 3960
DB 4871 CATGAGCTGACAGTGAAGCTGAGTGAAGCGGAGATTCGTTCTGAGATGAGATGAG 4930
QY 3961 ACCATGCGGATGCTGTTCCAGGAATGGAAGCAACCTGCTCCGAGCCTTGAAGTCT 4020
DB 4931 ACCATGCGGATGCTGTTCCAGGAATGGAAGCAACCTGCTCCGAGCCTTGAAGTCT 4990
QY 4021 CCGAGCTACCGGAGAGAGCTGTGAGAAAGGCTGAAAGCTTTGGCCAGCTCATCAAC 4080
DB 4991 CCGAGCTACCGGAGAGAGCTGTGAGAAAGGCTGAAAGCTTTGGCCAGCTCATCAAC 5050
QY 4081 AACAGGATGCTGCTGCTCTCAATCCGACAGCTGAGTCCAGAGTGTCTCCATG 4140
DB 5051 AACAGGATGCTGCTGCTCTCAATCCGACAGCTGAGTCCAGAGTGTCTCCATG 5110
QY 4141 CCGAGCTGAGCAAGTGTGCTCAATCATGACCGTGTGAGAGCAAGTGAAGTAC 4200
DB 5111 CCGAGCTGAGCAAGTGTGCTCAATCATGACCGTGTGAGAGCAAGTGAAGTAC 5170
QY 4201 GGCATGATGCTGAGAGAGCTGCTGGCGGAGCTCATTTGACAGAACTGAGAGCAAG 4260
DB 5171 GGCATGATGCTGAGAGAGCTGCTGGCGGAGCTCATTTGACAGAACTGAGAGCAAG 5230
QY 4261 AACCACTTAAGCTGCTCTGAGAGAGCTGAGTCAAGTGTGAGAGATGCTGACCAAT 4320

ORIGIN

Query Match 51.7%; Score 2939.6; DB 6; Length 6387;
Best Local Similarity 71.3%; Pred. No. 0;
Matches 3952; Conservative 0; Mismatches 1579; Indels 15; Gaps 5;

151 TTCAATCAGCTGTGTGATGAGAGACAGACATTTACTTGGGGGCGCTCAATCG 210
421 TTCAACACTTGAACCGTCCACCAAGGAGCGGGGCGGTCTATGTGGGGGCATCAACCG 480
211 ATTTCAGAGCTTCCAGCGACCTGAAGCTTGTGTGACGATGAGACAGGCGCGAGCG 270
481 GTCTATAGCTGACAGGCAACCTGACCATCAGGTGGCTCATTAAGACAGGGCCAGAAAG 540
271 GACACCCCAAGTGTATCCCAACCCGCGATGTCACAGCTGACATAGCCCTGACACCC 330
541 GACACAAAGCTTGTATCCCGCCCTCATGTGTGACGCTTGGACGGAAGTCTCACCTTC 600
331 ACCAACATGTCAACAAAGATGTCTCTCATATGACTTACAGAGAGAAAGGCTGATGCTGT 390
601 ACCAACATGTCAACAAAGATGTCTCTCATATGACTTACAGAGAAAGGCTGATGCTGT 660
391 GGGAGCGCTGTACCAAGGCACTGTGCAAGCTGTGAGGTGAGGAGCTCTTCAAGCTGGG 450
661 GGGAGCTTACCAAGGAGGCTGTGCAAGCTGTGAGGTGAGGAGCTCTTCAAGCTGGG 720
451 GAGCCTTATCATTAAGAGAGAGACATATCTGTGAGGTGCAACGAGAGCGGCTCACTTT 510
721 GAGCCTTATCATTAAGAGAGAGACATATCTGTGAGGTGCAACGAGAGCGGCTCACTTT 780
511 GAGATATGCTTCCCTACAGCAACCTGTGAGCAAGCTGTTCATTCGCCAGCGAGAGAT 570
781 GGGGTGATGTGTGAGCTGTGAGGTGAGGATGTGCAAGCTTCTTCACGCGTGTGAT 840
571 GGGAGAGCCGAGATATTTCCCAACATCTCCAGCGGAGAACTGACCAAGAACTGTGAGCG 630
841 GGGAGAGGATTAATCTCCGACCTGTGTCCAGCGGAGAGTGTCCCGAGACCTGTGATCC 900
631 GATGGCATGTTCGGTACGCTTCCATGATGATGTGTGTGCGCTCGATGATTAAGATCCT 690
901 TCAGCCATGTCTCATATGAGCTACACAGGATTTGTCTCTCTCTCATCAAGATCCCT 960
691 TCAGCACTTCAACATCATCCCTGACTTGTATATCTACTATGTATGTATTTAGCACT 750
961 TCAGCACTTCAACATCATCCCTGACTTGTATATCTACTATGTATGTATTTAGCACT 1020
751 GGGCACTTGTCTACTTTTGAACCTTCCAACTGAGAA--TGGTGTCTCCACAGGCTTC 807
1021 GGGGCGCTTGTCTACTTTCTCATCTGTCCAGCCGAGAACCTCGAGGGGTGTGGCATCAAC 1080
808 ACCACCAAGAGAGAGGTGTATACATCAAGCTGTGAGGCTTTGCAAGAGAGACAGGCG 867
1081 TCCGCTGAGAGACCTCTTCTACACTCAAGCATGTGTGGGCTGTGCAAGATGACCCCAAG 1140
868 TTCACCTCTATATGAGGTGCGCATGTGTGTGAGCGAGTGTGGGTGAGATGACGCGCTG 927
1141 TTCACCTCTATATGAGGTGCGCATGTGTGTGAGCGAGTGTGGGTGAGATGACGCGCTG 1200
928 CTGAGAGGCTGACATGCTGCAAGAGGCGGGGCGGTCTTGGAGAGACCTTGGAGGATCAT 987
1201 CTGAGAGGCTGACATGCTGCAAGAGGCGGGGCGGTCTTGGAGAGACCTTGGAGGATCAT 1260

LIILIVILIAVKRSRNDLTLKRLQMONDLBSRVALJECKEAPAEIOTDINETS
DIBRSGLPIYDIYTYAMRVLPFGIEDHPURELEEVANGQOHTKLPQIOLNNKV
FLIFLITRELOFSFWRDGNVASHLMTGLQRLIYADVLQILSLIDLNLEKN
HPLILRLTESVAEKLTWPAFLHFLKEGAPFLFMYCAIKQMKRGPIDATG
EAYSISEDLIRQOIEYKTLINCVNPDENBEPFLPVKLNCDITTOYKEHILAVY
KNVYSORPAVMDLEMRQRLARVLQEDIDITTKIEGKMLNTLMYQSDSVV
ALVKSOTSI PASAGISRTSISRYSRPRYGPDSLSRSPMTTPDLBSGKVMH
LVKNHDKDQKEDRGSKWSEYIYRLIATKGTLOKPDUDIPELTFSYHNGSALPL
AIKMPDLEQADBRSHIDTVRHTKSKCLPFRVWVYITPQVPIHSGSITDA
CLSVAGTENDSCSTSEHRLKDSPSKLIYADIPSYSWERYADIAKIPALSDO
DNMAYLAEQRLHAVERFNMALNEIYSYKSEELIGALBDEQARRQRLAYKEQ
LINAMSTES"

988 CCAGATGATACCTGCTCTTCAACCGCTCTTCTTCACAGGCGCAAGAGCCGAAATGAAATCC 1047
1261 AGCCAGAGACATGATCTCTTTTGCCATCTTCTCCAAAGGGGCAAGGATATCACACCCG 1320
1048 CTGATGAGTGGGCGCTGTGATCTTCACTTTGAAGCAGATTAATGACCGGATTAAGAG 1107
1321 CCCATGATCTGTCCCTGTGTGCTTCCCTATCCGGGCGATCAACTTGCAGATCAAGAG 1380
1108 CGGCTGAGCTTTGTTCACGGGGCGAGGGCAGCGTGAACCTGCGCTGATCAAGGTGAAG 1167
1381 CGCTGAGCTCTGTCTACCAAGGGCGAGGGCAACCTGAGGCTCAACTGCTGCTGGGAGAG 1440
1168 GACATCCCTGAGAGTGGGCTCTTAACATGACATGACATGACATGACATGACATGACATG 1227
1441 GACGTCCAGTACCAAGGCGCTGTCCCATCAATGATATCTTGTGACTGACATTC 1500
1228 AATGCTCCCTGGAGAGTGTCCGACATGTGTGGTGAATTCCTGCTTTCACGAGAGACAG 1287
1501 AACGAGCCCTGGAGGCTCAACTCAGTGAAGGGCTGACCTGTACACACAGCAGAG 1560
1288 GACGCGATGAGCTGTGTATGACATGATGCTTACAGAAACCATCTGTGCGCTTGTGGGC 1347
1561 GACGCGATGAGCTGTGTGCTCTTACGTTTACACGCTTACAGCGTGTGTGTGTGGGG 1620
1348 ACCAAAGTGGCAAGCTGAAAGATCCGGGTGATGAGACCCAGGGGCAACGCGCTCCAG 1407
1621 ACTAAGATGACAGCTGAAAGATTCGGGCGACAGGTGCCCGCCCATGATGGGTCCAG 1680
1408 TATGAGAGGTGACAGGTGT--GAGCCCGGCGCATCTCTCGGGATATGCTTCTCC 1464
1681 TACGAGAGGTGTGTGTGCTTACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1740
1465 AAGACCAAGAGCACTCTACATCATGTACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1524
1741 ATTGATGAGCGCTACCTGTATGATGATGTGAGAGAGAGAGAGAGAGAGAGAGAGAG 1800
1525 TCTGTGTGATGATCAGAGCTGCGGAGAGTGTGCTTGGCTACAGGCAACCCCATGTGGC 1584
1801 TCATGTGAGAGATATACAGCTTGTGGGAGAGTGTGAGAGAGAGAGAGAGAGAGAGAG 1860
1585 TGTGTGTGCTGCAACACAGTGTGACCCGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1644
1861 TGTGTGTGCTGCAACACATGTGTCTCCGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1920
1645 GCAAGTGTGCTGAGAGATGAGAGAGTGTGTCCGGTGAACGATCCATCCCAACATATC 1704
1921 AATGATTTGTGTCAGAGATCAGAGAGTGTGTGAGCTTGTGAGATCCAGCAGCATC 1980
1705 TCGTCTCTCAGTACAAAGT 1764
1981 TCAGTATGTGACACAGCGGTGTCTTACCTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 2040
1765 GCTGCGTCAACTGACCTTTGAGAGAGCTGTACAGATGATGAGTGTGTGTGTGTGTGTGT 1824
2041 GGGGTATGCGCTGT 2100
1825 CAGATCAGTGTCTCTGCGAGCGCAAGAGAGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1884
2101 CAGGTCATCTGACATCTACCTGCGGCGCAAGAGATGT--CCCTGTGATCCGCTGATCTAA 2157
1885 GACCAACATGTGTGACAGCTTCACTCAATCAAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1944
2158 GACTGT 2217
1945 ACCAGCTTGTCTTCAATGAGAGGTGTGACCAATTTGTGTGTGTGTGTGTGTGTGTGTGT 2004
2218 ACCAGTTCAGATTTTCAATGAGAGGTGTGACCAATTTGTGTGTGTGTGTGTGTGTGTGTGT 2277
2005 CCATACGCTGCTCACTGTGTGAATATCCGGATGTGTGACACCATGACCCCAAGAGCTGC 2064
2278 GCTTTCGCTGCTTGT 2337

QY	2065	TCCTTCCAGGAAGGCGGAGTGAAGTGGCCGAGGACCTGGCCCCAGGTCCTGCGAGTGGAC	2124
Db	2338	TCCTTCCAGGAAGGCGCGGATCAATATTTCAGAGACTGTCCCCAGCTGTGTCCACAGAG	2337
QY	2125	AAGATCCTGTGTCCCTGTGAAGGTGATCAAGCCCTATCAGCTGAAGGCCAAGACTTCCTCC	2184
Db	2338	GAGATCTTGATTCCAGTGTGGGGAGGTAAAGCCATCACCTTTAAGGCGCGAATCTGTGCC	2457
QY	2185	CAGCCCCAGTCTGGGACAGGTGGCTACGAATGCACTCTCCAACTTTCAGGGCAGGAGAG	2244
Db	2458	CAGCCGCAAGTCCGGCCAGCGAGGCTATAGTGTGTCTCAACTTACAGAGGCCATCTCAC	2517
QY	2245	CGAGTCCCGCCCTTGCGCTTCAACAGCTCCAGCGTCAAGTCCAGAAACCTCTTATTC	2304
Db	2518	CGGGTCCCCGCTCTGGCGCTTCAACAGCTCCAGCGTCAAGTGTCAAGACAGTCTGTACAG	2577
QY	2305	TATGAAGGATGAGATCAACAACTGTCGCGGAGTTGAAGATGCAAGTGTGTGAATGGGAC	2364
Db	2578	TATGATGGCAATGGACATCAAGCAATGTGGCCGTGATTTTCGTGTGTGTGAAGCGCAAT	2637
QY	2385	TTCAACATTGACAACTCCAGCTCGAGATAAGTTCACTCTTCAAGTGTGAAGCAATGGCT	2424
Db	2638	TTCAATCATTTGACAACTCCAGGACCTGTAAGATCATCTCTCAAGTGTGAGCCAGCGG	2697
QY	2425	GAGAGCTGCGGGCTGTGCTCAAGGCTGACCCAGACTTCGATGCGGTGTGTGCAAGGCG	2484
Db	2638	GAGAGCTGCGGCTCTGCTCTCAAGGCCAACCGGAAGTTGAGTGTGTGTGTGACAGCGC	2757
QY	2485	CCAGGCGAGTGCACCTTGGCCGACGACTGCTCTGCCCCAGGAGGCGAGTGGCTGGAAGT	2544
Db	2758	GAGGCGAAGTGCACCTTCCACGACACTGTACAGCCCTTCACGCCCCCTGTGCTGACATGG	2817
QY	2545	TCTGGTGGCCAAAGCAAGTGCACAAACCCCGCATCAGAGATTAATCCCGTCAAGGC	2604
Db	2818	TTCAAGCCAAATGTCAATGTCTCCAACTCCAAATCACGGAATTTTGAAGGTGTCTGGA	2877
QY	2605	CCCCGGGAGGGGGCAACCAAGTCACTATCCGAGGGGGAACCTGGGCTTGGAAATTTGCG	2664
Db	2878	CGCGCGGAGGAGGGGACCGAGTGCACATTCATGCGGGAACCTGGGCTTGGAACTTCTCC	2937
QY	2665	GACATCGGCTCCCATGTCAAGGTTGCTGAGCTGACAGCCCTTATGTAAGTGTATAC	2724
Db	2938	GAGATCGGCCACATGTGTGCAAGTGTGGGTGGGTCGCCCCCTCCAGGGGAATAC	2937
QY	2755	ATCCTCGAAGAACAGATGTGTGTGAGATGGGGGAGGCGCAAGCCGACAGCATGCAAGC	2784
Db	2998	ATCATCGTGTAGCAAGATGTCTGTAGATGGGCCATGCCCCGTGTGGGAACCACTTCGGG	3057
QY	2785	TTTCGTGAGATCTGTGTGAGTGTGTGTGTGGGCTGAAATTCATGSCCGGCTCTCAAGCTC	2844
Db	3058	CGAGTAGCCCTGTGTATTTGGGAGTGTGAAGCAAGATTCAAGAGAAATCCCATCAAGAG	3117
QY	2845	TATTACTTCATGACACTGACTCTCTCAAGATCTGAAGCCGACCGGGGGCCATGTCCGGA	2904
Db	3118	TACACTTTCGTGAACCTTCTGTGTGTGTGTCACTCAACCCAAATTCAGAGTCCCGAATCGGA	3177
QY	2905	GGAACCAAGTGCACATCAACAGGACCAACAACCTGAATGCGGAGCAAGCAAGTGTGTATG	2964
Db	3178	GSCATTAATGTGTACCATTAACGGGCCATTACCTTGGGGGTGGGAAGAGCTGTGGACTTAC	3237
QY	2965	TTTGGAAAAGCAGCCCTGTGTCTCTTCCACAGGGGATTCATCTCTATTTGTCTG--CAAC	3021
Db	3238	CTGGGCAACCAAGACTGTGAGTTCTAAGGAGGTCAATGAAGTGAATGTGTGTCTCA	3297
QY	3022	ACCAATCTTCAGATGAGGTGTCTAAGATGAAGGTGTGGTGCAGGTGGAACAGGCGCAAG	3081
Db	3298	CCCCCATATCAATGAGGCTTGGCCCGGTCCTGTTTCTGTGAAGTGTGACCGAGCCAT	3357
QY	3082	ATCCACCAAGACCTTGTCTTTTCAGATGTGTGGAAGCCCCACCAATGTGTGGGATTTGAGCA	3141
Db	3358	GTGGATGAACACTTCAGTGTGTGATGACAAATGATGACCTTCGGGTTCACAGGATGTGAAGCA	3417
QY	3142	GAATGAGCATTTGTAGTGAAGCAACCCATTCGCGGTATGGGGGACCCACCTTGAACCTC	3201

Db	3418	GAGTGGAGCATTGGCAATGGGCACACACCCCTTACATCAAGGCTTCAACCTGGATGTC	3477
QY	3202	ATACAGAAACCCCAAGTCCGTGCCAAGCATGGAGGAAAGACATCATATCTGGAG	3261
Db	3478	ATTCAAGAGGCCMAAGATCCAGTCAATTCAATGGCAAAAGATCTGCAATGTTGTA	3537
QY	3262	GTTCGAAAGCTATGATGATGACTCTGTACAGGCGCCCGCTCGCTGGGTCCTGACAC	3321
Db	3538	GTTGTGAACAACAACCCCTCACCCTGGCACCTCTGTACCAAGCATACCCGCT	3597
QY	3322	CAGTCAGACCTGACCCGAGAGGCGCGAGAGTTTGCTTCACTCTGGAACAATCTCATC	3381
Db	3598	GAGCTTGACACTGTGGAAAGCCCAAGATGATTTGATTTGCTTTAAAGTCCATATCC	3657
QY	3382	CTGCTCATTCCTCCAAACAAGACCACTTCACTACTATCCCAACCCGAGTTTGAGCCCTT	3441
Db	3658	TTGTGTAATTTCACACGACACCAAGTTTATCTACTACCCCAACCCGACTTTGAATGCTCT	3717
QY	3442	GGTCCCTCAGGAATCCTTGGAGCTTAAGCTTGGACGCCCTCATCTCTTAAAGGCAAAAC	3501
Db	3718	AGCCCTACTGAAAGCTTGGATCAAAACCCAGAGATGCCCATCATCTGAAAGGCAAAAC	3777
QY	3502	CTGATCCCGCCTGTGGCTGGGGGCMAAGTGAAGCTGAACATACTGTGCTGTGGGAG	3561
Db	3778	CTGAGCCCTCTGCTCTGGAGG--GGCCAAATCAACTACATGTGCTCATGGAGAG	3834
QY	3582	AAGCCCTGGACCGTGAACCGTGTCAAGATGTCAAGTGTCTGGCAGTCCCAACTCATC	3621
Db	3835	ACCCCTGTGTGTGACCGGTATCTGAACCCAGTTCCTTGGCAGGCTCCCACTTCAACC	3894
QY	3622	GGCAGGCAACAAGTATGAGCCCGCTGTGCGTGGCATGAGTACTCCCGGAGATGTTATAC	3681
Db	3895	GGGACAGCAAAAGTCTTGTTCACTGATGGGCGGATGTGTCTGCTGCTCGCTCGTATGAT	3954
QY	3682	ATTGCCCCCGACACCCCGGTCAAGCTTCCCGGCACTGTCAGCATCCGAGATGGAGTGGCGGC	3741
Db	3955	GTCAATTCAGACACTTGTGACCTTCCACAGCATGTGACATGGCGCGCGCGGAC	4014
QY	3742	CTCTCATCATTTTTCATCGTGGCCGTGCTCATTTGCTATAAAGCAAGTCCCGGAAAGT	3801
Db	4015	CTCTCTCATCATCGTCAATCATCATGTCTCTCATTTGCTTACAAGCGCAAGTCTCGAAGAAAT	4074
QY	3802	GACCTCAACGTGAAGCGGCTGCAATGACATGAGCAACAAGTGAAGTCCCGTGGCCCTG	3861
Db	4075	GACCTCACTCCAAAGCGGCTGCAATGACATGAGCAAACTGAGATCCCGTGGCCCTG	4134
QY	3862	GAGTGCAGAGAGCTTTGGCGACCTGCAAGCGACATCATGACTGACCAATGACTG	3921
Db	4135	GAGTGCAGAGAGCTTTGGCGACCTGCAAGCGATATCATGATGATTGACCAAGTGAAGCTG	4194
QY	3922	GATGAGCGCGGAGATTTCGTTCTCTGAGCATATGAACCTTACAACAGCGGAGTCTTCCCA	3981
Db	4195	GACCGCTTAGAATCTCTTACCTTGGACTATGTACTGTAAGCTATGCGATGCAAGTCTTCCG	4254
QY	3982	GGAATTGAAGCAACCTGTCTCCGCGGACCTTGAGTCCCGGCTTACCGGAGAGCGT	4041
Db	4255	GGCATTCGAGGACCAACCCGCTCTCCGGGAGCTGAGGCTTACAAGGAAACCGGCAACACAC	4314
QY	4042	GTCGAGAAAGGCTGGAAGCTCTTGCGCAAGTATCATCAACAACAAGAGTTCCTGCTGCC	4101
Db	4315	GTCGAGAAAGGCTTGAAGCTCTTGGCCAGCTCATCAACAACAAGGTTCTCTCTGAC	4374
QY	4102	TTCAATCCGACGCTTGAATGCCAGCTGATGCTTCTCATGCGGCAACGCTGGCAAGTGGCC	4161
Db	4375	TTCAATCCGACCTTGGAGCTGCAACGCAAGTTCTCTCATGCGGCAACCGGCGCAAGTGGCT	4434
QY	4162	TCATCATCATGATCGAGCTGCAAGCAAGTGTGAGTACGCACTGATGTGTCTAAGCAG	4221
Db	4435	TGCTCTCATGATCGGCTGCAAGGCGCGCTGGAATATGCACTGATGTCTCTAAGAG	4494
QY	4222	CTGCTGCGGACTTCATTTGACAGAAGCTGAGAGCAAGAACCAACCTTAAGTGTCTCTC	4281

Db 4495 CTGCTCTGTGACCTCATGATTAAGAACCTGGAGAACCAACCCCAAGCTGCTATC 4554
 QY 4282 AGGAGACTGATGATGAGTGTGAGAGATGCTGACCAATGGTTTACTTTCCTCTCTAC 4341
 Db 4555 CGGAGGACAGAGTGTGTGGCTGAAAAAGATGTCACAAATGGTTTGGCTTCTCTGAC 4614
 QY 4342 AAGTTCTCAAGAGTGTGTGGAGAGCCCTCTTCTCTCTGTTCTGTGCTATGACGAG 4401
 Db 4515 AAGTTCTCAAGAGTGTGTGGAGAGCCCTCTTCTCTCTGTTCTGTGCTATGACGAG 4674
 QY 4402 CAGATGAG 4461
 Db 4675 CAGATGAG 4734
 QY 4462 GACAACTCATGCGGACAGAGATGACTAAGAACCTGTGCTGAGCTGTGTACGCA 4521
 Db 4735 GACAACTCATGCGGACAGAGATGACTAAGAACCTGTGCTGAGCTGTGTACGCA 4794
 QY 4522 GACAACTCATGCGGACAGAGATGACTAAGAACCTGTGCTGAGCTGTGTACGCA 4581
 Db 4795 GACAACTCATGCGGACAGAGATGACTAAGAACCTGTGCTGAGCTGTGTACGCA 4854
 QY 4582 GTCAGAGAGAGAGATGCTGATGCTGATGCTGATGCTGATGCTGATGCTGATGCTG 4641
 Db 4855 GTCAGAGAGAGAGATGCTGATGCTGATGCTGATGCTGATGCTGATGCTGATGCTG 4914
 QY 4642 GCTCAAGATGATGATGCTGATGCTGATGCTGATGCTGATGCTGATGCTGATGCTG 4701
 Db 4915 GCAATGACATGATGCTGATGCTGATGCTGATGCTGATGCTGATGCTGATGCTGATG 4974
 QY 4702 GAAACATGACATGCTGATGCTGATGCTGATGCTGATGCTGATGCTGATGCTGATG 4761
 Db 4975 GAGACATGACATGCTGATGCTGATGCTGATGCTGATGCTGATGCTGATGCTGATG 5034
 QY 4762 CAGGTGCTGATGCTGATGCTGATGCTGATGCTGATGCTGATGCTGATGCTGATG 4821
 Db 5035 CAGGTGCTGATGCTGATGCTGATGCTGATGCTGATGCTGATGCTGATGCTGATG 5094
 QY 4822 GTGAAACATGCTGATGCTGATGCTGATGCTGATGCTGATGCTGATGCTGATGCTG 4881
 Db 5095 CTTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 5154
 QY 4882 ACCGGGACGCGGACGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 4941
 Db 5155 ACCGGGACGCGGACGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 5214
 QY 4942 GGAATCAAGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 5001
 Db 5215 GGGGTCAAGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 5274
 QY 5002 CGGGGAGAGAGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 5061
 Db 5275 CGGGGAGAGAGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 5334
 QY 5062 CTGAGAGAGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 5121
 Db 5335 CTGAGAGAGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 5394
 QY 5122 GCCCTGCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 5181
 Db 5395 GCTCTGCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 5454
 QY 5182 GGCATTATGACCGGACGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 5241
 Db 5455 AGCATTCATGACAGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 5514
 QY 5442 TGGGTCAACATGATCAAGAACCGGACGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 5301
 Db 5515 TGGGTCAACATGATCAAGAACCGGACGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 5574
 QY 5302 GAGCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 5361
 Db 5575 GAGCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 5634

QY 5362 CGGCTGGGCAAGAGACTCGCCCTCCCAAGAGCTGCTGATGCTGCAAGAGATCCCGAGCTAC 5421
 Db 5635 CGGCTGGGCAAGAGACTCGCCCTCCCAAGAGCTGCTGATGCTGCAAGAGATCCCGAGCTAC 5694
 QY 5422 AAGATTTGGGTGAGAGATTTATCTCAGACATPAGGAGAGATGCCAGCATGAGAGACAA 5481
 Db 5695 AAGAGCTGGGTGAGAGATTTACTAGCAGACATCCGCAAGCTCCCAAGCATPAGTACGAG 5754
 QY 5482 GACATGAACGATACCTGCTGAGAGAGCTCCCGATGACATGAATGACTTCAACACCATG 5541
 Db 5755 GACATGAATGCTACTCTGCGGAGAGAGCTCCCGCTGACAGCGCGTGAAGTTCAATGCTG 5814
 QY 5542 AGTGACCTCTAGAGATCTTCTCTATGTGGGCAATACACGAGAGATCTTGTGACT 5601
 Db 5815 AGTGCCCTCAATGATGATCTACTCTATGTGCAAGATGATGAGAGCTCATCGGGGCG 5874
 QY 5602 CTGACACAGATGACATGCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 5661
 Db 5875 CTGAGAGAGATGAG 5934
 QY 5662 ACCCTCATGAGCTTACAGAGCTGAAA 5687
 Db 5935 AATGCCATGCTCATGAGAGCTGAGA 5960

RESULT 5
 AB007932 6263 bp mRNA linear PRI 13-AUG-1998
 LOCUS Homo sapiens mRNA for KIAA0463 protein, partial cds.
 DEFINITION
 AB007932.1 GI:3413887
 VERSION
 KIAA0463 protein.
 KEYWORDS
 Homo sapiens (human)
 SOURCE
 Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 ORGANISM
 1 (sites)
 REFERENCE
 AUTHORS
 Seki, N., Ohira, M., Nagase, T., Ishikawa, K., Miyajima, N.,
 Nakajima, D., Nomura, N. and Ohara, O.
 TITLE
 Characterization of cDNA clones in size-fractionated cDNA libraries
 from human brain
 JOURNAL
 DNA Res. 4 (5), 345-349 (1997)
 MEDLINE
 PUBMED
 9455484
 REFERENCE
 2 (bases 1 to 6263)
 AUTHORS
 Ohara, O.
 TITLE
 Direct Submision
 JOURNAL
 Submitted (08-OCT-1997) Osamu Ohara, Kazusa DNA Research Institute,
 Laboratory of DNA Technology, Yama 1532-3, Kisarazu, Chiba
 292-0812, Japan (E-mail:cdna@info.kazusa.or.jp, Tel: +81-438-52-3913,
 Fax: +81-438-52-3914)
 FEATURES
 source
 1..6263
 Location/Qualifiers
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /chromosome="1"
 /clone="HG0942"
 /sex="male"
 /tissue_type="brain"
 /clone_id="pbluescriptII SK plus"
 1..6263
 /gene="KIAA0463"
 <1..5894
 /gene="KIAA0463"
 /codon_start=3
 /product="KIAA0463 protein"
 /protein_id="BAA3308.1"
 /db_xref="GI:3413887"
 /translation="GGGRLPDSLELRAGSRRPQQAALVDAAGTGGCTDAVLES
 CEGQKAVQMSHRSRLTAPLSMRKRPALVDSVYLLSVYLLAPRA
 GMPQSTSHSENRMWTFNHLTVHQGTAVVGAALNRYKLTGNLTIOVAHRTGPEBDN

KACVPLIYQPCSEVLTLLNNVKKLLIDYSENRLLAGSLYGVYCKLRLDILIV
 EBSHKEHYLSVNTKGTWYIVASBEGDKLFTGIVADQKDFPPLSRKPRDP
 ESSAMLDYELSHDFVSLIKIPSDILALSHEDIFYIIGFASGGVYVLPVOPFPG
 VAINSAGDLYFYSIRIVRLCKDPEKHSVYSLFPFGCTRGVRRLLQAAVYLAKEPSSLA
 QAFNITSODVLFALFSPKQKQYHHPDDISALCAFPALINQIGRIQSCYQGBNL
 OLVNLTGKQVOCFAVPIDNFCGLIDNPLGASFPVGLTYTTSRDMTSVASVY
 VNGSVYFVFGTSKRLKIRADGPBGVGYEMVSVLQKSLIDMPSIDORLIVY
 NSBROTVRPVPSCEQYTCGBCISGDPGCMCALHMCSSRSDCOOAMPNRAAS
 ISOCVLAHPSSISISESRILSLVSNAPPLSAGIACBENLVEVGSVYVIC
 ISGEPDVPIPLDQWFGELQKSKETGKI FSTERFETGKALHOLCSNAPR
 CHWCKYRNLCTHDPITCSFQBERINISECPQVLPBELLFVGAVKILTKARLPQ
 POSGQREYECVNLIGAIHVRPALRNSSVQCONSVOYDGMDSINAVFAVWNG
 NFIIIDNPOLKYLKCAORSCGLCLADRFECGMSGERCTDLOHCSFSPSP
 LDMSSNAINSONPOTIELTVSGPSPGRTVTHGNLCLDSELAHVQVAGCTP
 LPEVYIAOVCEMHALVGTSGPVRCLCECKBEHMTSHOQYTVSNVSLNP
 TRPESGGMVTITGHYLNGASSVAVYLIDNCEPFGMSITVCSPPSSNGSPV
 VSVSVRAHVDNLOPEYIDDPVOKIIEBMSIASGHTPLITGNRLVQPRRVK
 FNKESVNVCKVNTTTLTCLAPSLTDDPGLDVERDEFGVFNNOVSLIYNDT
 KFIYVNPTEFELSLPTGLDOKRGPBIIKGNLCPBPASGAKMYVLIGETPCAVT
 VSETOLCEPPLTGQKMYVGVKVPSPSVYSIDSLTLPALIVISIAAGSILLI
 IYIYVIAVRKRENDLTLEKLOMOMDLBSRVALCEKEAPAELODINELTSLDR
 SGIPLYDYRTVAMRVLPGLIEDHPVLRELVGNGOQHEKALKELOINNKVILT
 FITTELOQSFMSKRDGNVSLIMTGOURLBYADYKQILSDILDNLEKHPKX
 LIRTESVAEKMLTNFAFLHFKPEKCEBEPFMYLCATKQOMKGPIDATVNAV
 YSOPRAVMDLEMRQRIARVVLQDEDITKIEGKWLNTIMHYQVDSRVVALVP
 KORSYNIIPASASISRTSISRYDSFRYGPSPLSRAPMTPLDESQKWLHVN
 HDHDOKEGRSGKNSVSEIYELTALKTLOKVDLEETLESYVHRSALPILIKY
 MPDPLEOGRDHSIHDTDVHRTKNSVCLPFRMVNTIKPOFVPIHGSITDALCY
 VAOTPDSCSTSEHRLGKSPSKLIXADIDISYKVRVRYADIAKIPALSDOMNA
 YLAEQRLHAHVEFNMLALNELIYSVSKISELIGLEDOGRORRAVYKQELINA
 MSIES"

ORIGIN

Query Match: 51.6%; Score 2938; DB 9; Length 6263;

Best Local Similarity 71.2%; Pred. No. 0;

Matches 3951; Conservative 0; Mismatches 1580; Indels 15; Gaps 5;

151 TTCAATCAGCTGTGTGATGAGAGACAGACATTTCCTGGGGGCGCTCATCGG 210
 Db TTCAACACCTTACCGTCCACCAAGGAGCGGGGCGCTCTATGTGGGGGCGATCAACCGG 416
 357 211 ATTACAGCTCCCGACGACCTGAAGCTGTGTAGCGATGAGACAGGGCGGAGCG 270
 Db 417 GTCTATAGCTGAGAGGCAACTGACCATCAGGTGCTCTATAAGACAGGGCCAGAAAG 476
 477 271 GACAAACCCAAAGTGTACCAACCCCGCATGCTCCAGACCTGCAATGAGCCCTGACACC 330
 Db 477 GACAAACAGTCTGTACCCGCCCTCATGTCGACAGCCCTGACAGCAAGTGTCAACCTTC 536
 331 ACCAACAATGTCACAAAGATGCTCCATAGACTACACAAAGAAAGAGCGCTGATTCCTGT 390
 Db 537 ACCAACAATGTCAACAAAGCTGCTCATATGACTACTCTAGAAACCGCTGCTGCGCTGT 596
 391 GGGAGCCCTGTACCAAGGATCTGCAGCTGTGAGGCTGAGAGACCTCTTCAAGCTGGG 450
 Db 597 GGGAGCCCTCTACCAAGGGGCTGTGCACTGTGCGGCTGAGATGACCTTCACTCGTGG 656
 451 GAGCTTATCATTAAGAGAGACATATCTGTCAAGTGTCAACAGAGAGCGGCTCAAGTCTT 510
 Db 657 GAGCATTCACCAAGAAAGAGACACTACTGTCCAGTGTCAACAGACGAGGACCAAGTAC 716
 511 GAGAGTATGCTGTCTCTACAGCAACTGTGATGACAAAGTGTTCATTGCGACGAGATGGAT 570
 Db 717 GGGGTATGTGGCCTCTGAGAGGTGAGAGATGCAAGTCTTCACTGCGACGAGCTGGAT 776
 571 GGGAAAGCCGAGTATTTTCCACCATCTCCAGCCGAAACTGACCAAGAACTCTGAGGCG 630
 Db 777 GGGAAAGCAGATTAATCTTCCACACCTGTCCAGCCGAAAGCTGCCCGAGACCTTGATTC 836
 631 GATGCAATGTCGAGTACGCTCTTCATGATGATGATTCGTGAGCTGATGATTAAGATCCCT 690
 Db 837 TCAGCAATGCTGACTATAGACTACACAGGATTTTGTCTCTCTCACTAAGATCCCT 896

691 TCAGCAATGCTGACTATAGACTACACAGGATTTTGTCTCTCTCACTAAGATCCCT 750
 Db TCAGCAATGCTGACTATAGACTACACAGGATTTTGTCTCTCTCACTAAGATCCCT 956
 897 751 GGGCAATTTGTACTTTTGAACCTTCAACCTGAGAA--TGTGTCTCCACAGGCTCC 807
 Db 957 GGGGAGCTTTGTACTTTTCACTGTCACCTGACCGAGACCCCTAGAGGTGTGGCATCAAC 1016
 808 ACCAACAAGAGCAGTGTATACATCCAACTGTGTGAGGCTTGAAGAGAGACAGCC 867
 Db 1017 TCGCTGAGACCTCTTCACTCACTCAAGATGTGTGGCTCTGCAAGATGACCCCAAG 1076
 868 TTCAACTCTATGATGAGGTGCCCATTTGCTGTGAGCGAGTGGGTGAGTACCGCTG 927
 Db 1077 TTCAACTCATACGTGTCCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1136
 928 CTGAGAGCTGCTTACCTGTCCAAAGCGGGGCGGTGCTTGGCAGAACCTTGGAGTCCAT 987
 Db 1137 CTGAGAGCTGCTTACCTGTCCAAAGCGGGGCGGTGCTTGGCAGAACCTTGGAGTCCAT 1196
 988 CCAATGATGACCTGCTCTTACCGCTTCTCCAGGGCCAGAAACGAAATGAATTC 1047
 Db 1197 AGCCAGAGCATGATGATCTTGTGCTGCTTCTCCAAAGGAGAGAGATGATCAACCGC 1256
 1048 CTGATGATGAGGCTGCTGCTGCTGCTTCTTCAAGAGCATTAATGACCGCATTAAGAG 1107
 Db 1257 CCGCATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1316
 1108 CCGGTGAGTGTGTATACCGGGGCGAGAGGACAGCTGAGCTGAGCTCAAGAGTGAAG 1167
 Db 1317 CCGGTGAGTGTGTATACCGGGGCGAGAGGACAGCTGAGCTGAGCTCAAGAGTGAAG 1376
 1168 GACATCCCTGAGCAGAGTGGCTTCTTACCATGATGATGATGATGATGATGATGATGATG 1227
 Db 1377 GACGTCCAGTGTGACCAAGAGGCGCTGCTCCATGATGATGATGATGATGATGATGATG 1436
 1228 AATGCTCCCGGAGAGTGGCAGCATGATGATGATGATGATGATGATGATGATGATGATG 1287
 Db 1437 AACGAGCTCCGAGAGCTCACTGAGTGAAGGCGCTGAGCTGATGATGATGATGATGATG 1496
 1288 GACCGCATGAGCTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1347
 Db 1497 GACCGCATGAGCTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1556
 1348 ACCAAAGTGGCAGAGTGAAGAGATCGGGGTGATGAGACCAAGGAGCAAGCGCTCCAG 1407
 Db 1557 ACTAAGATGGCAGAGTGAAGAGATCGGGGTGATGAGACCAAGGAGCAAGCGCTCCAG 1616
 1408 TATGAGAGGATGAGAGTGT--GAGCCCGGCGAGTCTCGGAGATGATGATGATGATGATG 1464
 Db 1617 TATGAGAGGATGAGAGTGT--GAGCCCGGCGAGTCTCGGAGATGATGATGATGATGATG 1676
 1465 AAGAGCAGAGCAACTTATCATGATGATGATGATGATGATGATGATGATGATGATGATG 1524
 Db 1677 AATGATGAGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1736
 1525 TCTGTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1584
 Db 1737 TCTGTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1796
 1585 TGTGTGTGTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1644
 Db 1797 TGTGTGTGTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1856
 1645 CCGAGATTTGCTCGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1704
 Db 1857 AATGATTTGTCTGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1916
 1705 TCGGTCTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1764
 Db 1917 TCAGTATGAGCAAGCGGCTTGTGATGATGATGATGATGATGATGATGATGATGATGATGATG 1976

QY 1765 GCTGCGTCACTGACCTTTGAGAGCTGTGAGAGATGAGATGGCTGGCTGGGCAAT 1824
Db 1977 GCGGATATGCGCTGTGCTTTGGGAACTGACAGAGATGAGAGGAGGTGTGCCGAGC 2036
QY 1825 CAGATCAAGTGTACTTCCCTGCAAGCAAGAGAGTCCCGGATCTACAGAGATGGG 1884
Db 2037 CAGGTATCTGCACTCTCAGCTGGGCCAAGAGATGT---CCGTGTCTCCGCTGATCA 2093
QY 1885 GACCAACATGCTGTACAGCTTCACTCAATCAAGAGAGAGGAGCTGACTCCGAGC 1944
Db 2094 GACTGTGTTGGGCTGAGAGCTACAGCTGAGGTCCAGAGAGAGAGATTTGTGACG 2153
QY 1945 ACCAGCTTGTCTTCAAAATTGCAAGGTCCACATTTGCTGCTGTCTGCTGAGAGT 2004
Db 2154 ACCAGATTCAAGTTTAAACATGAGTCCAGTCCCACTGTGCTGTGTCTGTCAACG 2213
QY 2005 CCATACCGCTGCACTGTGTAAATACCGGATGTGTGACCCATGACCCCAAGACTG 2064
Db 2214 GCTTCCGCTGCTGATGTGTGCAAGTACCGCACTTGTGCACTGTATCCCACTGCTG 2273
QY 2065 TCCCTTCAAGAGAGGCGGAGTGAAGCTGCCGAGAGTGCCTCCAGCTGTGCGAGTGA 2124
Db 2274 TCCCTTCAAGAGAGGCGGAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGA 2333
QY 2125 AAGATCTGTGCTGCTGAGAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGA 2184
Db 2334 AAGATCTGTGCTGCTGAGAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGA 2393
QY 2185 CAGCCCAAGTCTGTGAGAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGT 2244
Db 2394 CAGCCCAAGTCTGTGAGAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGT 2453
QY 2245 CAGTGTGCTGCTGCTGAGAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGA 2304
Db 2454 CAGTGTGCTGCTGCTGAGAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGA 2513
QY 2305 TATGAGAGAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGA 2364
Db 2514 TATGAGAGAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGA 2573
QY 2365 TTCAACATGAGCAACCAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGT 2424
Db 2574 TTCAACATGAGCAACCAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGT 2633
QY 2425 GAGAGCTGCGGCTGTGCTGCAAGGCTGCAAGGCTGCAAGGCTGCAAGGCTGCAAG 2484
Db 2634 GAGAGCTGCGGCTGTGCTGCAAGGCTGCAAGGCTGCAAGGCTGCAAGGCTGCAAG 2693
QY 2485 CCAAGGCTGCAACCTGCGCAGCACTGCGCAGCACTGCGCAGCACTGCGCAGCTG 2544
Db 2694 CCAAGGCTGCAACCTGCGCAGCACTGCGCAGCACTGCGCAGCACTGCGCAGCTG 2753
QY 2545 TCTGTGCTGCAAGGCTGCAAGGCTGCAAGGCTGCAAGGCTGCAAGGCTGCAAGG 2604
Db 2754 TCTGTGCTGCAAGGCTGCAAGGCTGCAAGGCTGCAAGGCTGCAAGGCTGCAAGG 2813
QY 2605 CCGCGGAGAGGAGGCAACAGTCACTATCCAGAGGAGAACTGAGGCTGGAATTCG 2664
Db 2814 CCGCGGAGAGGAGGCAACAGTCACTATCCAGAGGAGAACTGAGGCTGGAATTCG 2873
QY 2665 GACATGCTGCTGCTGCAAGGCTGCAAGGCTGCAAGGCTGCAAGGCTGCAAGGCTG 2724
Db 2874 GACATGCTGCTGCTGCAAGGCTGCAAGGCTGCAAGGCTGCAAGGCTGCAAGGCTG 2933
QY 2725 ATCCCTGCAAGAGATGCTGTGTGAGATGAGAGGAGGCAAGGCTGCAAGGCTGCAAG 2784
Db 2934 ATCCCTGCAAGAGATGCTGTGTGAGATGAGAGGAGGCAAGGCTGCAAGGCTGCAAG 2993
QY 2785 TTGTGAGATGCTGT 2844
Db 2994 TTGTGAGATGCTGT 3053
QY 2845 TATTAATGACATGACATGCTCTCTGATCTGAGATCTGAGCCAGCGGAGGCTGCTGCA 2904

Db 3054 TACACCTTCCGAACTCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 3113
QY 2905 GGGAGCCAGAGTCACTACAGGAGCAACCTGATGAGGAGGAGGAGGAGGAGGAGGAG 2964
Db 3114 GGGAGCCAGAGTCACTACAGGAGCAACCTGATGAGGAGGAGGAGGAGGAGGAGGAG 3173
QY 2965 TTGTGAGATGCTGT 3021
Db 3174 TTGTGAGATGCTGT 3233
QY 3022 ACACATCTGATGAGT 3081
Db 3234 ACACATCTGATGAGT 3293
QY 3082 ATCCAGAGAGCTGT 3141
Db 3294 ATCCAGAGAGCTGT 3353
QY 3142 GAATGAGATGCTGT 3201
Db 3354 GAATGAGATGCTGT 3413
QY 3202 ATACAGAACTCCAGATCCGAGCAAGTGAAGGAGGAGGAGGAGGAGGAGGAGGAG 3261
Db 3414 ATACAGAACTCCAGATCCGAGCAAGTGAAGGAGGAGGAGGAGGAGGAGGAGGAG 3473
QY 3262 GTTGTGAACGCTACTGATGATGATGATGATGATGATGATGATGATGATGATGATG 3321
Db 3474 GTTGTGAACGCTACTGATGATGATGATGATGATGATGATGATGATGATGATGATG 3533
QY 3322 CAGTGAACCTGAGACCGAGAGGCGGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 3381
Db 3534 CAGTGAACCTGAGACCGAGAGGCGGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 3593
QY 3382 CTGCTATCTTCAACAGACCACTTCACTATCTTCACTATCTTCACTATCTTCACTAT 3441
Db 3594 CTGCTATCTTCAACAGACCACTTCACTATCTTCACTATCTTCACTATCTTCACTAT 3653
QY 3442 GGTGCTGAGAACTGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 3501
Db 3654 GGTGCTGAGAACTGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 3713
QY 3502 CTGATCCGCTGTGCTGTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 3561
Db 3714 CTGATCCGCTGTGCTGTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 3770
QY 3562 AAGCTGAGACCTGAGACCTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGT 3621
Db 3771 AAGCTGAGACCTGAGACCTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGT 3830
QY 3622 GGGAGGCAAGAGTGTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 3681
Db 3831 GGGAGGCAAGAGTGTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 3890
QY 3682 ATTGCGCGGAGACCGCTGAGCTGTGAGCTGTGAGCTGTGAGCTGTGAGCTGTGAGCT 3741
Db 3891 ATTGCGCGGAGACCGCTGAGCTGTGAGCTGTGAGCTGTGAGCTGTGAGCTGTGAGCT 3950
QY 3742 CTGCTATCTTCAACAGACCACTTCACTATCTTCACTATCTTCACTATCTTCACTATCT 3801
Db 3951 CTGCTATCTTCAACAGACCACTTCACTATCTTCACTATCTTCACTATCTTCACTATCT 4010
QY 3802 GACCTGAGCTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 3861
Db 4011 GACCTGAGCTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 4070
QY 3862 GAGTGAAGAGAGCTTGTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 3921
Db 4071 GAGTGAAGAGAGCTTGTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 4130
QY 3922 GATGAGCGGAGATTCGCTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAG 3981

QY 137 AGCCCCCGAGGGTTTAAATCACTGTGTGTGATGAGAGACAGACATTTACTTGG 196
Db 683 ATGTGTGATGGAGTTTCAACATTTGTACTGTACACGAGAAACAGAGGGCTGTGTATGTGG 742
QY 197 GGGCCGTCATTCGATTTTCAAGGCTCCAGCGACCTGAAAGTCTTGTGTGACGATGAGA 256
Db 743 GGGCTTCAATTCGTGTTCAGAGTGTGACGTGCAACCTCAACATCCAGGTGTCTCAACA 802
QY 257 CAGGGCCGAGCAGAGACACCCCAAGTTTACCCACCCCGCATGTCTCCAGACTGCAATG 316
Db 803 CAGGGCCAGAAAGAGACAAAGAGCTTGTCTACCCACCCCTCATTTGTACAGCCCTGCAATG 862
QY 317 AGCCCCGACACACACAAATGTGTCAAAAGATGTCTCATAGACTACAGAGAGACA 376
Db 863 AAGTGTTCATCTACCAACAAATGTCAACAACTAGTATCATTTGACTACTGTGAATATC 922
QY 377 GGGTGAATGCTGTGGAGGCTGTACCAAGGCATCTTGCAAGCTGTGAGGCTGAGAGCC 436
Db 923 GCGTGTGGCTGTGGAGGCTGTACCAAGGCATCTTGCAAGCTGTGAGGCTGAGAGCC 982
QY 437 TCTTCAGAGTGGGAGGAGCTTATCATAGAGAGAGAGCATCTGTGAGGTCAGAGAGA 496
Db 983 TCTTCATCTGTGTGAGGAGCTTCCACAGAGAGAGAGCATCTGTGAGGTCAGAGAGA 1042
QY 497 GCGGCTCAGTCTTGTGAGTATGCTCTCTACAGCAACCTGATGACAAAGCTGTTCATG 556
Db 1043 CAGGCAACATGTATGTGTATGTGTGTGGCTCTGAGGGGAGAAATGGCAAGCTTTATG 1102
QY 557 CCAGGAGATGTATGGAGAGCCGAGATTTTCCACATCTCCAGCCGAGAACTGACCA 616
Db 1103 GCACTGTGTGTGAGAGAGAGATTACTTCCCTACTGTGTCAAGCCGAGACCTGCCCC 1162
QY 617 AGAAGCTGAGAGGAGATGATGTCGCTACGCTGTCCATGATGATGATGATGATGATG 676
Db 1163 GTGACCTGTGATCTTTCAGCAATGTGTGATGATGATGATGATGATGATGATGATG 1222
QY 677 TGATTAAGATCCCTTGGACACCTTACCAATCATCTCTGATGATGATGATGATGATG 736
Db 1223 TCATCAAGATTCCTCTGACACCTTACGCTGTCTCTCTCTCTCTCTCTCTCTCTCTCT 1282
QY 737 ATGTTTTAGAGAGGCACTTGTCTACTTTTGAACCTTCAACCTGAGATGATGATGATG 796
Db 1283 ATGCTTTGCTGAGTGGGGGTTTGTCTACTTTCTCACTGTCCAGCCAGACACCCCTGAG 1342
QY 797 CACCAAGGCTCA---CAACAGAGAGAGATGATATCATCAAGCTGTGAGGCTTTGCA 853
Db 1343 GCATGGCCATTAATTCAGCTGAGAGACCTTCTATACCTCAAGATTTGTGCTCTGCA 1402
QY 854 AGGAGACACAGCCTTCAACCTGATGATGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 913
Db 1403 AGGATGACCCCAAGTTCACCTCTATGTGTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 1462
QY 914 TGAATACCGGCTGTGAGGCTCTACCTGTCTCAAAAGCGAGGAGGAGGAGGAGGAGGAG 973
Db 1463 TGAATATCGGCTTCTGAGGAGCACTTACCTTGCAAGGAGGAGGAGGAGGAGGAGGAG 1522
QY 974 CCGTTGAGATTCATCCAGATGATGATGATGATGATGATGATGATGATGATGATGATG 1033
Db 1523 CCGTTCAACATGAGAGGAGAGAGATGCTGTGTGCTGTGTGTGTGTGTGTGTGTGTGTGT 1582
QY 1034 GGAATATGAATCCCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1093
Db 1583 AGTACACACACCCCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1642
QY 1094 ACCGATTAAGAGAGGCTGTGATGATGATGATGATGATGATGATGATGATGATGATG 1153
Db 1643 TGAATATCAAGAGAGGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1702
QY 1154 GGGCTCAAGGATGAAGACATCCCTGAGAGAGATGAGGCTTTAAACATGAGATGATGATG 1213
Db 1703 GGGCTGTGAGAGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1762
QY 1214 GTGAGCTGAGACATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1273

Db 1763 GGGCTGTGAGACATCAACAGGCTTGGAGAGGCTCACTCCGTGTGAGAGAGCTGACCTGT 1822
QY 1274 TCACGAGAGACAGAGACCGCATGACGTCTGTCAATGATGATGTCTACAGAGACATCTC 1333
Db 1823 ATACCAAGACAGAGAGACCGCTTACCTGTGTGAGCTCTCTATGTTTAAATGCTTACAGTG 1882
QY 1334 TGGCTTTGTGAGGACACAAAGTGTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1393
Db 1883 TGGTTTGTGAGGACATGAGTGTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1942
QY 1394 GCAAGCCCTCCAGTATGAGACGAGTGAAGTGT---GAGCCCGGAGCTCTCCGAG 1450
Db 1943 ATGATGAGGATCAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 2002
QY 1451 ATATGAGCTTTCACAGAGACACAGAGACATCTCTACATCATGTGAGAGAGAGAGAGAG 1510
Db 2003 ACATGAGCTTTCATCATATGATGATGATGATGATGATGATGATGATGATGATGATGATG 2062
QY 1511 GAGTCCCTGTGAGTCTGTGTGATGATGATGATGATGATGATGATGATGATGATGATGATG 1570
Db 2063 GGGTCCCTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 2122
QY 1571 ACCCCACTGT 1630
Db 2123 ATCTCAGCTGT 2182
QY 1631 GGTCCAGAGAGCCCGGAGGTTGT 1690
Db 2183 GGGCTGTGAGAGACATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 2242
QY 1691 ATCCCAACATATCTCCGTCTCTGATGATGATGATGATGATGATGATGATGATGATGATG 1750
Db 2243 ACCCCACAGATCTCTGT 2302
QY 1751 TCCCGAGCTGT 1810
Db 2303 CTCCCAACCTCTGTGAGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 2362
QY 1811 TGT 1870
Db 2363 AGATATGTGAGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 2419
QY 1871 TCACGAGATGAGAGACACCATGATGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1930
Db 2420 TCCCTGT 2479
QY 1931 TGACCTTCCGACGACACAGCTTGTCTTCAATGAGAGGTCACAAATTCGTGTGTGTGT 1990
Db 2480 AGATCTTGT 2539
QY 1991 CTTGT 2050
Db 2540 CTTGT 2599
QY 2051 ACCCCAGAGCTGTCTCTTCCAGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 2110
Db 2600 ACCCCAGATCTGTCTCTTCCAGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 2659
QY 2111 TGT 2170
Db 2660 TGT 2719
QY 2171 CCAAGAGCTCCCGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 2230
Db 2720 CCGAGAGCTCCCGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 2779
QY 2231 AGGAGAGGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 2290
Db 2780 AAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 2839
QY 2291 ACACTTTATTCATGAGAGGATGAGATCAACCTGAGGAGGAGGAGGAGGAGGAGGAGGAG 2350

Db	2840	ACACGCTGACACGATGATGATGGGATGAGCATTCAGGAACTGAGAGCTGGACCTTTCGTGTAG	2893
Qy	2351	TGTGGAAATGGGCACTTTCACATTTGACAACTCCAGCTCAGATTAATAGTTCACTCTTAACAAGT	2410
Db	2900	TATGGAAATGGGCACTTCACTTAATTTGACAACTCCAGGACCTGAAAGATCATCTCTACAGT	2959
Qy	2411	GTGAGACCATCGTGTAGAGCTGCGGGCTGTGCTCTCAAGAGCTGACCCCACTTCCCATGTG	2470
Db	2960	GTGCAAGCCCAACCGGGAAAGCTGTGTGTCTCTGCTCTAAGGCTGACCAACAAGTTCGAGTGTG	3019
Qy	2471	GCTGTGTCGAGGGCCCGAGGCCAGTGCACCTGTCCGACGACTGTCCCTGTGCCAGAGAGCC	2530
Db	3020	GCTGTGTGAGTGTGTGAGGCGAGATGTACCTCTCCACACGACCTGTCCCGACACTTCTGTAGCC	3079
Qy	2531	AGTGGCTGAGACTGTCTGTGTGCTCCAAAGGCAAGTGCACAAACCCCGGACTCAGAGATTA	2590
Db	3080	CCTGGCTTGACTGTGTCCAGGCCAACATGTCAAGTGTTCACACCCCAATATCAGAGATTT	3139
Qy	2551	TCCGGGTACAGGCCCCCGGGAGGGGGACACAAAGTCACTATCCGAGGGGAAACTGTG	2650
Db	3140	TGAACATATCAGGACCACTTAAGAGGGAGCTGTGTGATCCATCATGTGCGTGAACCTGTG	3199
Qy	2651	GCTGTGAATTTCCGAGCAATCCGCTCCCATGTCAAGGTTGCTGTGGGTGAGTGCACCTT	2710
Db	3200	GCTTGGACTTCTGTGAGTGTCTACCATGTGAGGTGGCTGTGAGTGTGCGCTGCACACTTA	3259
Qy	2711	TAGTGTGATGTTCATCCCTGTGACAAACAGATCTGTGTGTGAGATTTGGGGAGGGCCAAAGCCA	2770
Db	3260	TCCAGGGGAAATACATCATCTGCTGACGAGATGTGTGTGAAATAGGCGCATGCGCTTAATAG	3319
Qy	2771	GCCAGCATGACAGGCTTCCGTGAGATCTGCGGTGCTGTGTGTGTGCGGCTGAAATTCAATGCGCC	2830
Db	3320	GTACCACTCTGTGGCGCTGTGGCGCTGTGCAATTTGGGGAAATGCAAGCAGAGTTCATGACCA	3379
Qy	2831	GGTCTTCACACTCTATTACTTCATGACACTGTCTCTCAAGATCTGAAAGCCGACCGGG	2890
Db	3380	AGTCCCAACACAGTATACCTTTGTGAACTCTTGTGTGTCTCACTCAGCCCGGATCCGGG	3439
Qy	2891	GGCCCATGTCCGAGGGGACCCAGTGCACATCACAGGCAACCACTGAATGCGGGAAACA	2950
Db	3440	GACCAAGATCAGGAGGTACATGTGTGACCATCACAGGCCATTACTTGTGTCTGGAGACA	3499
Qy	2951	ACGTGTGTGTGATGTTTGTGGAAGGACCGCTGTCTTCCACAGCGATCTCCATCTTACA	3010
Db	3500	GTTGTGCAGTGTACTCTGGGCATATCAGACTGTGAATTTATGTGGAGGTCAATGATGTGGA	3559
Qy	3011	TTGTCTGCACACA--CCACATCTTCAGATGAGGTCTTAAGATGAAAGGTGTGCGTGCAGG	3067
Db	3560	TTGTATGTGTTCACCCCATTCATTCATTCAGTACGTAGACCAAGTCCCTGTCTCCGTGAGTG	3619
Qy	3068	TGGAAGAGGGCCAAAGTTCACACAGACCTGTGTCTTCAATGTGTGGAACACCCCACTATCG	3127
Db	3620	TGCAAGAGGCCCGGGTGTGTAGCACTGTGTGCACTTGTGATATATAGATATCCACGGGTCC	3679
Qy	3128	TGCGGATTTAGCCAGCAATGAGCACTTGTCAGTGAACAACCCATCGCGTATGGGGGA	3187
Db	3680	AACGATTTAGGCAAGAGTGTGATATCACTAATGGGGACACACCCCTTAACATCACAGGCT	3739
Qy	3188	CCCAAGCTGAACTCTATACAGAAACCCCGAGATCCGTGGCAACAGTGAAGAGGAGAGACA	3247
Db	3740	TCACTTTGATGTCTATTACAGAGCCCAAGGTTCAGTCAAAATTATGTATGSCAAAGATCTG	3799
Qy	3248	TCAATATCTGTGAGGTTCTGAAACGCTACTGAGATGACTGTTCAGAGCGCGCGCTCGCTC	3307
Db	3800	TCAATGTATGACAGTGTGTAAACAACAACACCCCTCACTGTCTGTGCACCTCTCTGTACCA	3859
Qy	3308	TGGGTCTGACCAACAGTCAAGCTGACCGAGAGGCCCGGAGAGTTTGGCTTCACTCCGTGG	3367
Db	3860	GTGACTACCGCCCAAGTCTGTGACACTGTGTGGAAGGSCAGATGAGTTTGTGATTTCTCTTTA	3919
Qy	3368	ACAACGTCCAGTCTCTGTCTCATCTCTCAACAAGACCAACTTCACTTATCTCCAAACCGG	3427
Db	3920	ACAATGTTCAATCTTACTCATCTTATACAGACACCAAGTTCATCTACTACCCCAACCCA	3979

QY	3428	TGTTGAGGCTTTGGTTCCTCAGGAATCTCGAGCTCAAGCTCGGACGCCCATCATCC	3487
Db	3980	CGTTTGAACGTCTACGCCCCCATGGAATCTTGATCAAGAGCCAGGCTCACCCATCATCC	4039
QY	3488	TAAAGGCAACACCTATCCCGCTTGCGTGGGGCAACGTGAAGCTGAACCTACACTG	3547
Db	4040	TGAAGGGCAAAAATCTGTGCTCTCTCTCTCTGGAG---GGCCAACTCACTACAG	4096
QY	3548	TGCTGTTGGGGAACAGCCGTCGACCGCTGACCGCTGACATGTCACACTGCTCGGAGT	3607
Db	4097	TAAATGTTGGAAGAACCTTGACAGTCACTGTGTGTGAGACACAACTGCTTTGTGAAC	4156
QY	3608	CCCCCACTCATCGGACGACGACCAAAATGATGCCCCTGTCCGTTGCGATGAGTACTCC	3667
Db	4157	CTCCCAACTCACAAGGAGACCAAGGTCAATGTTCACTGAGGGGGATGGTGTTCAC	4216
QY	3668	CGGGATGATGTATATTTGCCCGGACAGCCCGCTCACCTGCCCGCATGTCAGACTCG	3727
Db	4217	CTGGCTCCGTAGGCTCATCTTCGACAGCCTGTTGACCTTCGACGATCATCAGCATCG	4276
QY	3728	CAGTGGCTGGCGGCTCTCATCATTTTCATCGTGGCCGTGCTCATTTGCTATTAAAGCA	3787
Db	4277	CAGCTGGGGAAGCCTCCTTCTTATCTGTCATCATGTCTTCATGCTTACAAAGGCA	4336
QY	3788	AGTCCCGGAAGGACCTCACTCGTGAAGGGCTCGAGATGACATGACATGACCACTGGAGT	3847
Db	4337	AGCTTAGGAATAATGACTTCACTCAAGCGGCTTCCAAATGCAATGGAACACTGGAGT	4396
QY	3848	CCCGTGTGGCCCTGAGTGCAGGAAGCAAGCACTTTGCGAGCTGCAGACGCAATCCATGAC	3907
Db	4397	CAAGGATGCACTGAGTGCAGGAACCTTTTGGGAGCTTCAACAGACATCATATAGC	4456
QY	3908	TGACCAAGGACATGATGAGAGCGGGGATTCGGTTCCTGGACATATAGAATTACACATGC	3967
Db	4457	TAAACAGTGAATTGATGATGATCAGAAATCCCTTACTCGGACTACCGTATACCTATGCTATGA	4516
QY	3968	GAGTGTCTTTCCAGGAATTGAAGACCACTGTCTCTCCGGACCTTGAGGTCCGGGCT	4027
Db	4517	GAGTCTCTTCCAGGACATTGAGGACCACTGTCTCTCCGGGAACGTGAGGTACAGGGAA	4576
QY	4028	AACCGGAGGACGATGTGGAGAAAGGCTGAAAGCTCTCGCCACATCAACAACAAG	4087
Db	4577	ATGGACACACACGTTGGAGAAACCTTGAAACTCTTGCCGACTTATCAACAACAAG	4636
QY	4088	TGTTCTGCTGTCTTCAATCGACGCTTGAAGTCCAGCTTGAAGTTTCAATGCGCAAC	4147
Db	4637	TGTTCTGCTGACCTTCAATCGTCACTGTGAATCAAGCGGACGCTTCTCATATGGAAC	4696
QY	4148	GTGGGAACGTGGCTCACTCATATGACCGTGCAGAGCAAGTGAAGTACCCCACTG	4207
Db	4697	GTGGGAACGTGGCTCTTCAATCAATGACAGGCTTCAAGGTCGCTTGAATATGCCACTG	4756
QY	4208	ATGTGCTGAACAGCTGTGGCGCACTCATTTGAAGAACCTGGAAGACAGAACCAAC	4267
Db	4757	ATGTCTCAACAGCTGCTCTCTGACCTCATTTGAAGAACCTGGAAGACAGAACCAAC	4816
QY	4268	CTAAGCTGTCTCAGAGGACATGATGATGAGTGTGAGGAATGATCTGACCAATTGGTTTA	4327
Db	4817	CCAACTGTCTCTCGGACAGACATGATGTGTGGCGGAGAAATGCTGATCACTGCTTTG	4876
QY	4328	CTTTCCTCTTCAAGATTCTCAAGAGATGTGTGGGAGCCCTCTTCTCCTGTTC	4387
Db	4877	CTTTCCTCTTCAACAAGTTCTCTGAAGAGTGTGTGGGGAACAACCTCTTCACTGTAATCT	4936
QY	4388	GGGCATCAAGAGAGATGAGAGAGGCCCCCATGACGCATCAAGGGCAGAGCCGCT	4447
Db	4937	GTGCATTAAGACGATGAGAAAAAGCCCCCATTGAAGCTATTAATGATGTGAGGCCGAT	4996
QY	4448	ACTCCTTAGCGAAGCAAGCTCATTCGCGCACAGATTAAGTACTACAAAACCTGTCTCTGA	4507
Db	4997	ACTCCTTAGTGAAGCAAGCTCATTCGCGCACAGATTAAGTATTAAGCTGTGATCTCTGA	5056

QY	4508	GCTGTGTGAGCCAGACAACTATCCAAACAGCCCGAGAGTCCAGATGAAGATCTTCAACTGTG	4587
Db	5057	ACTGTGTCAACCTCGACATATGAACAGCCAGAGATCCAGATGAAGTACTTAACATGTG	5116
QY	4568	ACACCATCTACTCAGGTCAGAGAAAGATCTTGATGTCATCTTCAAGATGTGCTTGTCT	4627
Db	5117	ACACCATCTACTCAAGTCAGAGAAAGATCTTGATGTCGATATGAAGATGTCTCCACT	5176
QY	4628	CCCAACCGGCCAAAGCTSCAGATATGATCTTGAGTGGCCGACAAAGAAAGTGGGGCAAGGA	4687
Db	5177	CCAGCGGCCAAAGGCTTGGACATGATCTTGAGTGGCCGCAAGGCCGGAATTGCCAG	5236
QY	4688	TGATCTTGACAGATGAAGACATCCACCAAGATTTAGATATGATTTGAAGCGACTGAACA	4747
Db	5237	TGTGTGTTCAAGAGCAAGACATTTACCAAAATATGAGGGTGACTGGAAGCGGCTTTAA	5296
QY	4748	CACGTGGCCACTACTACAGGTGCCAGATGTTCCGTGTGTGGCATTTAGTGTCCAGACAGTGA	4807
Db	5297	CACGTAGTCATTTACAGGTGTCCAGACATCCGTGTGTGGCTCTGTGTTCTTAAGACACT	5356
QY	4808	CAGCCTTAATACGACAGTGAACAATCCACCGCTCTCCAGGACCTCAGCAAGTAAATATGAAA	4867
Db	5357	CCTCTTAACAATCCGTGCTCTGTCCAGCATCTCTCCGACATCCATTAAGACATATGACT	5416
QY	4868	ACATGATCCGGTACAACGGGCAAGCCCCGACAGCTTCGCTCAAGCACTATATGATCACTC	4927
Db	5417	CTTCCCTTACGATTAACAAGGCAAGCCAGACACCTCCGGTCCCGGATCCCACTGATCACCC	5476
QY	4928	CTGACCTGGAAGTGGAGTCAAGATGTGGCACTAGTGAAGAACCAAGAGACGAGAGCC	4987
Db	5477	CAGACTTGAAGAGGAGTCAAGGTTTGCACTGTGTGAAGAAATCAGACATGTGTACC	5536
QY	4988	AGAAGAGGGGGACCGGGGAGCAAGATGTGTGAATCTTACCTGACCCGACTCCG	5047
Db	5537	AGAGAGAGGTGACCGGGGAGCAAAATGATGTCTGAGATCTATCTTGAACCGGGCTTAC	5596
QY	5048	CCACTAAGGGCACTGCGAAGATTTGTGTGATGACCTTTTGAACCATCTTCAAGCACGG	5107
Db	5597	CCACCAAGGGGACCTCGCAAAATTTGTGAGCACTGTTTGAACCTTGTTCAGACACTG	5656
QY	5108	CACACCGGTGCTTGCCCTGCCCCGTGGCCATCAAGTACATTTTGACTTCTCTGGATGAGC	5167
Db	5657	TGCACCGGGGATAGTCTCTCCCTTAGCAATCAAGTACATGTTTGAATTTCTCGATGAGC	5716
QY	5168	AGGCTGATTAACATGTGCATTTATGACCCGCAAGCTCCGCATACCTCGAAGACAAATGGC	5227
Db	5717	AGGCAAGACAGACAGATTCACACACACAGATGTGCGGCAACCTCGMAAAGCAACTGCC	5776
QY	5228	TGCCCTGAGGTTTGGGTCAACATGATCAAGAAACCGCAATTTGTGTTTCAATCCATA	5287
Db	5777	TTCCACTTTCGTTTCTGGGTGAATGTCAATCAAGAACCCCAATTTGTATTTGACATCCACA	5836
QY	5288	AGAACAGCATCAACAAGCCCTGCTCTCTGTGTGTGCTCAGACCTTCATGAGACTTGTCT	5347
Db	5837	AGGCAAGCATCAAGATGCTGCTCTCTGTGTGTAGCCCAAGCTTTATGACCTCTGT	5896
QY	5348	CCAGGTCAAGACACCGGCTGGGCAAGACTGGCCCTCCAAAGATGCTGTATATCCAG	5407
Db	5897	CCACATCAAGACCCGACTAGGCAAGACTCACTCCAAAGATGCTGTATATCCAG	5956
QY	5408	ACATCCCAAGTACAAAGATTTGGGTGAGAGTATTAATCCAGATAGAGGAAGTGCAG	5467
Db	5957	ATATCCCAATTAATGAAGATGGGTAGAAAATACTATGCAAGATTTGCCAAGCTCCAG	6016
QY	5468	CCATACGCGACCAAGACATGAAGCATATCTGTGGCTGAGCACTCCCGATGCAATGAATG	5527
Db	6017	CCATTTAGCAAAAGATATGATGCTATCTCTCGGGAGCAAGTCCCGCTGATGCTATCAG	6076
QY	5528	AGTTCAACACCATGATGCACTTCAAGATCTTCTCTATGTGGGCAAAATACAGCAGG	5587
Db	6077	AGTTCAATATGTAGAGCCCTCAACAGATCTTACTCATATGTAGCAAGTACAGTGAAG	6136
QY	5588	AGATCTTTGACCTCTGGAACAAGATGACAGGTGTGGAGAGCAACTGCGCTTACAAAC	5647

[illegible]

[illegible]

ORIGIN

Query Match	48.8%	Score 2778.4;	DB 10;	Length 6730;
Best Local Similarity	69.3%	Pred. No. 0;		
Matches 3858; Conservative	0;	Mismatches 1691;	Indels 15;	Gaps 5;

QY	137	AGCCGCGGAGGTTTCATCACTCGGTGGTGAATGAGGACAGGACACATTTATCTGG	136
Db	683	ATCGTGACTGGACCTTTCACCACTTTGACTGTACTACACGAAACAGGGGGCTGTGTATGG	742
QY	197	GGGCGCTCAATCGGATTTTACAAGCTCTCCAGCGACCTGAAAGTCTTGTAACGATAGA	256
Db	743	GGGCTATCAATCGTGTCTTACAAGTTGACTGGCAACTTCACCATCGAGTGGCTCACAGA	802
QY	257	CAGGGCGCGACGAGAGCAACCCCAAGTTTACCCACCCCGATATGTCACACCTGCAATG	316
Db	803	CAGGGCCAGAAAGAGACAAAGAGCTTGCTACCAACCCCTATTGTACAGCCCTGCACTG	862
QY	317	AGCCCGTACCAACCAACCAATGTGCAACAAGATGCTCCATAGACTACACAGAGAGAA	376
Db	863	AAGTGTCTTACACTCACCAACATGTGCAACAACTACTGATATATGATCTACTGAAATTC	922
QY	377	GGCTGATTGCTGTGGAGAGCTGTACCAAGGCATCTGCAAGCTGCTGAGGTGAGAGCC	436
Db	923	GGCTGTGGCTGCGGAGAGCTCTTACCAAGGTGTTTGGAACTCTCTGCACTAATGATACC	982
QY	437	TCTTCAAGCTGGGGAGGCTTATATATGAAGAAGAGACATATCTGTACAGTGTCAACGAG	496
Db	983	TCTTCATCTGTGTGAGGCATTCACCAAGAAAGAAACATTACTTGTCCAGTGTCAATAGA	1042
QY	497	GCGGCTCAGTCTTTGAGATGATTCGCTCTCTACAGCAACCTGATGACAAAGCTGTTCATG	556
Db	1043	CAGGACCATGTATGTGTGTGATGTGGGCTCTGAGGGGAGAGATGCAAGCTTTTATACG	1102
QY	557	CCACGGCAGTGGATGGAGAGCCCGAGATTTTCCCAACATCTCCAGCCGGAACTGACA	616
Db	1103	GCACTGCTGTGGATGGCAAGCAGATTTACTTCCCTACTCTGTCCAGCGCGAAGTGCCCC	1162
QY	617	AGAACTCTGAGGCGGATGGCATGTTTGGGTAAGCTTCCATGATGAGTGTGGGCTCGA	676
Db	1163	GTGACCTCTGAAGCTTTCACCAATGTGGAACCTATGAGCTCCACAGATTTTGTCTCTCC	1222
QY	677	TGATTAGATCCCTTCGACACTCTTACACATCATCCCTGAATTGATATCTACTATGTCT	736
Db	1223	TGATCAAGATTCCTCTGTACACCCCTAGCCGTGTCTGTCACTGTGAACTCTTTCATATCT	1282
QY	737	ATGTTTATGACGTGGCAACTTTGTCTACTTTTGAACCTTCACCTGACCTGATATGTCTC	796
Db	1283	ATGGCTTTTGGCCAGTGGGGGGTTTGTCTACTTTCCTACTGTCACGTCAGACAGACCCCTGACG	1342
QY	797	CACCAAGGCTCCA---CAACCAAGSAGAGGTATATATCCAAAGCTGTGAGGCTTTGCA	853
Db	1343	GCATGGCCATCAATTCACTGAGACCTCTTCTTATACCTCAAGAAATTGTGTCTCTGCA	1402
QY	854	AGAGAGACACAGCCTTCAACTCTTATGTAAGAGTSCCAATTGGCTGTGAGCGCACTGGGG	913
Db	1403	AGGATGACCCCAAGTTTCACTCTATATGTGTCCCTGCTTTTGGCTGACACAGTGTGGGG	1462
QY	914	TGGAATACCGCTGTGTGCAAGGTGCTACTTGTCCAAAGCGGGGGCGTGTCTTGGCAGAG	973
Db	1463	TGGAATATGCGCTTCTGTGCAAGGAGCTTACTTGTCCAAAGCAGGGGAACTCTTAAGTCAGG	1522

QY	974	CCCTTGGAGTCCATCCAGATGATGACTGCTCTTCAACCCGCTTCTCCAAAGGCCAGAGC	1033
Db	1523	CCTTCAACATCAGACGACGAGATGATCTCTGTTGCAATCTTTCCAAAGGCGAGAGC	1582
QY	1034	GGAATAATAATCCCTGATGATGATCGGCCCTGTGCACTCTCATCTTGAGACGATAAATG	1093
Db	1583	AGTACACACACCCCTGATGATCTCTGCTCTGTGCTTCCCATCGGCGCATCAACT	1642
QY	1094	ACCCGATTAAAGACCGGCTCAGCTCTGTTAACCGGGCGAGGGGACGCTGACCTGGCCT	1153
Db	1643	TGCAATTCAGAGACGGTTGCACTCTCTCAACGAGAGGGCACTTGGACTCAACT	1702
QY	1154	GGCTCAAGGTGAGAGATCCCTCGACAGATGCGCTTTAACATTTAGCATATCTT	1213
Db	1703	GGCTGCTGGGAAAGATGTGAGAGCACAAAGGCGCTGTCCCATGAGATTACTTCT	1762
QY	1214	GTGGCCTGGACATGAATGCTCCCTGGGAGGTGCGACATGATGCTGTGAATTCCTCT	1273
Db	1763	GCGGCTGGACATCAACGACTTGTGGAGGCTCACTCTGTGAGAGGACTGACCTGT	1822
QY	1274	TCACGGAGGACAGGACCGCATGACGTCTGATCGATATGTCTTACAGAACCATCTTC	1333
Db	1823	ATACACACAGAGGAGACCGCTGACCTCTGTGGCTCTATGTTTAAATGGCTACGTG	1882
QY	1334	TGGCCTTTGTGGGACCAAAAGTGGCAAGCTGAAAGAAAGATCGGGTGGATGACCCAGG	1393
Db	1883	TGTTTTTTGTGGGACTTAAGATGAGTGGCAAGCTGAAAGAAATTCAGACGTATGTCCTCC	1942
QY	1394	GCAAAGCCCTTCAGTATGAGACGGTGGAGGTGT - -GACCCCGCCAGTCTTCGGG	1450
Db	1943	ATGATGGGGTCCATGATGAGATGCTCTGTGTTCAAAAGTGGAGGCCATCTTCGGG	2002
QY	1451	ATATGGCCTTCCAGAGCACAGACAGACACTCTCATCATCTGAGAGGACCTCACCA	1510
Db	2003	ACATGGCTTTCTTCATCAATAGCTATACCTATATGTCTCTGAGAGACAGGTACCA	2062
QY	1511	GAGTCCCTGTGAGTCTGTGTGTCAGTATCAAGACTGCGCGAGTGCCTTGGCTAGAGC	1570
Db	2063	GAGTCCCTGTGATCATGTGTAACAGATATCACTTTGTGAGAGATGTCTAAGCTAGGG	2122
QY	1571	ACCCCACTGTGGCGGTGTGTGCTGCACAAACATGTCACCCGGAAGGCGGTGAGC	1630
Db	2123	ATCTTCACTGTGTGTGTGTCCTGTCCCTGCACAAACATGTGTCCGAAGACAAATGCCAAC	2182
QY	1631	GATCCAGAGACCCCGACAGTTTGCTTGGAGATGAAGCAATGTGTCCGCTGACGTC	1690
Db	2183	GAGCTGTGGAAGCAATTCGATTTCTGTGCAATACAGCAATGATGATGAGCTTGAGTAC	2242
QY	1691	ATCCCAATATCTCGCTCTCAGTACAAACATGCTGCTGCTGCTCTGGAGACGTACATG	1750
Db	2243	ACCCCAACGATCTGTGTGTCAATCAACGCGGCTGTGAGCTGGTGTGATGATGATG	2302
QY	1751	TCCGGAAGCTGTAGCTGGCGCTCAATGCACTTTGAGACCTGTCAAGATGATGGGC	1810
Db	2303	CTCCCACTCTCTGAAGATATTCCTGTGTGCTTTGGAACTCTAATGAGGTGAGGGAC	2362
QY	1811	TGATGTGGGAATCAGATCAGTCCAGTCACTCCCTGACGACCAAGAGGTGCCCGATCA	1870
Db	2363	AGATATCTGGGAGTCAATCTATCTGCACTCACTGAGCCCAAGATGT - -CCCTGTCA	2419
QY	1871	TCACAGAGATGGGACCAACATGTCTGATCAGCTTCAACTCAATTCAAAGAGACCGGCA	1930
Db	2420	TCCCTCTGATCAACATGATGATTTGGCTTAAGCTGACAGCTAAGATCCAAAGACAGGA	2479
QY	1931	TGACCTTGACGACGACGACTTTGTCTTCAATTTGACAGGCTCACAAATTCGTCTGT	1990
Db	2480	AGATCTTTGTCAAGCACCGAATTCAAAGTCTATTACTGAGTGCCACCAACTGTGCTGT	2533
QY	1991	CTGTGCTGAGATCATACCGCTGTCCACTGTGTAAATTAACCGGATGTGCAACCCATG	2055
Db	2540	CTGTGTTTAAACGCGCTTCGCTGCAATGATGTGCAAGTACCGTAACTCTGTGACATGT	2599

QY 2051 ACCCCAGACCTGCTCTCTTCCAGAGAAAGCCGAGTGAAGTCCCGAGACTGCCCCCAGC 2110
Db 2600 ACCCCACTACCTGCTCTTCCAGAGAAAGCCGAGTGAAGTCTTCCAGAGACTGCCCCCAGC 2659
QY 2111 TGCTGGAAGTGAAGATCTGTGTCCTGTGAGTGAATCAAGCTTATCAGCTTAAG 2170
Db 2660 TGTGTCCACGAGAGATTTCTGATCCAGTTGGGAGTGAAGTGAAGTCAAGCTTAAG 2719
QY 2171 CCAGAACTCTCCCGCAGCCGAGTGTGGAGGCTGAGTCAAGTCACTTCAAGTTC 2230
Db 2720 CCGAAATCTGCCCCAGACCCCACTGTGCGCAGAGGCTTACAGTGTGTGCTCAGATTC 2779
QY 2231 AGGCGAGCGAGCGAGAGTGTGCGCTGTGCTTCAACAGCTCAGAGTCAAGTCCAGA 2290
Db 2780 AAGGCGCTGTCCACCGGCTCTGCTGCTTCAACAGTTCAGTGTGCAAGTGTCCAA 2839
QY 2291 ACACTCTTATTCCTTGAAGAGATGAGATCAACAACCTGCGCGTGAAGTTCAGAGTGC 2350
Db 2840 ACAGCTGTACAGATATGATGGATGGAATCAGCAACCTTACAGAGTGTGCTGTGAG 2899
QY 2351 TGTGAATGGGCACTTCAACATTTGACAAACCACTGAGTGAATTAAGTTCACTTCAAGT 2410
Db 2900 TATGAAATGGCACTTCAATTTATGACAAACCTGAGACTGAAGTACATCTTCAAGT 2959
QY 2411 GTGAGGCACTGCTGAGAGAGTGTGCGCTGTGCTTCAAGCTGAGTCCAGAGTGTG 2470
Db 2960 GTGAGGCGAGCGGAGAGTGTGTGCTGTGCTTCAAGCTGAGTCCAGAGTGTGAGTGTG 3019
QY 2471 GCTGTGTCCAGGCGCCAGGCAAGTGCACCTGTGCGCAGCAGTCCCTGCGCAGAGAGCC 2530
Db 3020 GCTGTGTCAATGTGTGAGCGAGATGTACCTTCCACAGCAGCTGCCAGAGACTTCAAGC 3079
QY 2531 AGTGTGTGAGCTGTGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2590
Db 3080 CTTGTGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 3139
QY 2591 TCCGCTGTGAGAGGCGCCGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 2650
Db 3140 TGAAGATATGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 3199
QY 2651 GCTGTGAATTTGCGGAGCATGCTCCCTCAATGTCAAGTGTGCTGCTGCTGCTGCTGCT 2710
Db 3200 GCTGTGCTTCTGTGAATTTGCTCAACATGTGAGGAGGAGGAGGAGGAGGAGGAGGAGG 3259
QY 2711 TAGTGAATGTATCATCTTCAAGAGAGATGCTGTGAGTGAAGGAGGAGGAGGAGGAGG 2770
Db 3260 TCCAGGAGGAGTATCATCTGCTGAGAGATGCTGTGAGTGAAGGAGGAGGAGGAGGAGG 3319
QY 2771 GCCAGATGAGGCTTGTGAGAGATCTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCT 2830
Db 3320 GTACCAATCTGAGGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGT 3379
QY 2831 GGTGTCAAGCTTATTTATTTATTTATTTATTTATTTATTTATTTATTTATTTATTTAT 2890
Db 3380 AGTGTCAACGAGATATTTATTTATTTATTTATTTATTTATTTATTTATTTATTTATTT 3439
QY 2891 GGCCCATGTCCGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 2950
Db 3440 GAGCAGAGTCAAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 3499
QY 2951 AGGTGTGTGTATTTTGTGAAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 3010
Db 3500 GTGTGGAAGTGTACTGTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 3559
QY 3011 TTGTCTGCAACA---CCACATCTCAAGATGAGTGTGAGATGAGATGAGTGTGAGTGTGAG 3067
Db 3560 TTGTATGTGTTCACCCCATTCATTCATTCATTCATTCATTCATTCATTCATTCATTCAT 3619
QY 3068 TGAAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 3127
Db 3620 TCGACAGAGGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 3679
QY 3128 TGGGATGTGAGCGAATGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 3187

Db 3680 AACGATTTGAGCCGAGAGGAGGAGTATCTAATGTGGGAGCACACCCCTTACATTCAGAGCT 3739
QY 3188 CCCACCTTGACCTATACAGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 3247
Db 3740 TCAATTTGATGTATTCAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 3799
QY 3248 TCAATTTGATGTATTCAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 3307
Db 3800 TCAATTTGATGTATTCAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 3859
QY 3308 TGGGCTTGAACACAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 3367
Db 3860 GTGACTACCGCCAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 3919
QY 3368 ACAAGTTCAGTCTCTGCTATCTCTCAACAAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 3427
Db 3920 ACAATTTCAATCTTACTATCTATTAACAGACACCAAGTTCATCTACTACCCCAACCA 3979
QY 3428 TGTTTGAGGCTTGTGTCTCTCAGAGATCTGTGAGTCAAGCTTGGCAGCGCCATCATCC 3487
Db 3980 CGTTTGAATGTCTAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 4039
QY 3488 TAAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 3547
Db 4040 TGAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 4096
QY 3548 TGTGTGTGGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 3607
Db 4097 TAATATTTGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 4156
QY 3608 CCCCCACCTGATGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 3667
Db 4157 CTCCACCTTACAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 4216
QY 3668 CGGAGATGTATTAATGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 3727
Db 4217 CTGCTCTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 4276
QY 3728 CAGTGTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 3787
Db 4277 CAGCTGTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 4336
QY 3788 AGTCCCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 3847
Db 4337 AGTCAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 4396
QY 3848 CCGGTGTGCTGTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 3907
Db 4397 CGAGGTGTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 4456
QY 3908 TGAACAGTGAATGTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 3967
Db 4457 TTAACAGTGAATGTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 4516
QY 3968 GGTGTGTGTCCAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 4027
Db 4517 GAGTCTGTGTCCAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 4576
QY 4028 ACCGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 4087
Db 4577 ATGAGCAGCAGCAGTGTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 4636
QY 4088 TGTTCCTGTCTCTTCAATCCGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 4147
Db 4637 TGTTCCTGTCTCTTCAATCCGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 4696
QY 4148 GTGAGCAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 4207
Db 4697 GTGAGCAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 4756
QY 4208 ATGTCTGAAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 4267

Db	4755	ATGTCCTCAACAGAGTGTCTCTCTACCTCACTTGAACAAGAACCTG93AGAACAAAGAACCAACC	4815
Qy	4268	CTAAGGCTGCTCTCAAGAGAGACTGAGTCAAGTGGCTGAGAAATGCTGACCAATTGGTTTAA	4327
Db	4817	CCAAAGCTGCTTCTCCGACAGGACTAAGTCTTGCGCCGAAAGATGTGACTAAGCTGGTTTG	4876
Qy	4328	CTTTCCTCCTCTTACAAAGTTCTCTCAAGAGAGTGTGCTGGGAGCCCTCTCTCTCCCTGTTCT	4387
Db	4877	CTTTTCTTCTTCAACAAGTTCTCTGAAGAGTGTGGCTGGGGAAACACTCTTCAATGCTTAATCT	4936
Qy	4448	ACTCCTTTGAGGAGAGCAAGCTCACTCCGCGACAGATTTGACCTAACAACCCCTGGTCTGA	4507
Db	4997	ACTCCCTGATGGAAGACAAGCTCACTCCGCGACAGATGAGATTAAGACTCTGATCTCTGA	5056
Qy	4508	GCTGTGTACGCGCCAGACAAATGCCAACAAGCCCGAGGTGCCGATTAAGATCCTCAACTGTG	4567
Db	5057	ACTGTGTAACTCTGACATAGAAACAAGCCCAAGATCCAGATGAAGTACTTAACCTGTG	5116
Qy	4568	ACACCATCACTCAAGTCAAGAGAAAGATTCTGATGCCATCTTCAAGATGTGCTTGTCT	4627
Db	5117	ACACCATCACTCAAGTCAAGAGAAAGATCCTCGATGCCGTATTATGAATGTCCCTACT	5176
Qy	4628	CCCAACCGGCCCCAAGCTGCACATTTGATCTGTGAGTGGCCGACAGGAAGTGGGCAAGGA	4687
Db	5177	CCCAAGCGCCAAAGGCTGTGACATGTGATCTGTGAGTGGCCGACAGGACCGGATTTGCCAAG	5236
Qy	4688	TGATCTTTGACAGATGAAAGACATCAACCAAGATTGAGATGATTTGAAAGCGACTGAACA	4747
Db	5237	TGGTGTTCAGAGACGAAACATTTACCACTAAATAGAGGCTGACTGGAAGCGGCTTTAA	5296
Qy	4748	CACGTGGCCACTACACAGSTGCGCAATGTGTCGGTGGCAATTGATGTCCAAAGACGTGA	4807
Db	5297	CACGTAGTGCATTTACAGSTGTGACAGATCCGATGGTGGCTGTGTTCTTAAGCAAGCT	5356
Qy	4808	CAGCCTTAATACGAGTGAACAATCCACCGTCTCCAGGACCTCGACAAATGAATATGAAA	4867
Db	5357	CCTCCTTAACAATCCCTGCTCTGTGCAACATCTCTCGACATCTCATTTGACAGATGACT	5416
Qy	4868	ACATGATCCGGTAAACAACGGGAGCCCCGACAGCTCCGTCACCGGCAACTATGATCACTC	4927
Db	5417	CTTCTCTTAGGTAACAAGGACGCCACAGACGCTTCGGTCCCGGTCCTCCATGATCAACC	5476
Qy	4928	CTGACCTTGAGAGTGTGACTCAAGATGTGGCACTTAGTGAAGAACCAAGACACGAGACC	4987
Db	5477	CAGACTTGGAAGCGGTGTCAAGGTTTGTGGCACTGTGTGAAGAAATCATACATGATGATGACC	5536
Qy	4988	AGAAAGGAGGGGGAGCCGGGGGAGCAAGATGTGTCTGAATATTAACCTGACCGGACTCCGAG	5047
Db	5537	AGAAAGGAGGGTGAACCGGGGACAGCAAAATGTGTCTGAATATTAACCTGACCGGCTTCAG	5596
Qy	5048	CAACTAAGGAGCACTGCAGAAAGTTTGTGATGACCTCTTTGAGACATCTTCAAGACGG	5107
Db	5597	CCACCAAGGAGCACTGCAGAAATTTGTGAGCACTTGTGAGACCTGTGTTCAGACCTG	5656
Qy	5108	CACACCGTGGCTCTGCGCTGCCCCCGGCACTCAATGATGTTTACTCTCGATGAGC	5167
Db	5657	TGCACCGGGGTAGTGTCTCTCCCCCTAGCCATCAAGTAAATTAATTTCTGTGATGAGC	5716
Qy	5168	AGGCTGATTAACATGGCAATTCAAGACCGGCACTGCGCCCATTACTCTGAAAGACAAATTGCC	5227
Db	5717	AGGCAAGACAGACACAGTATCCACGACACAGATGTGCGGACACACTGGAAAAAGCAACTGCC	5767
Qy	5228	TGCCCCCTAGGTTTGGGTTCACATGATCAAGAACCCGCAAGTTTGTGTTGACATCATTA	5287
Db	5777	TTCCACTTCGTTTCTGGGTGATGTGATCAAGAAACCTCAATTTGTATTGACATCCACA	5836
Qy	5288	AGAACGACATCAAGACGCTGCTCTCTGTGTGTGGCTTCAAGCTTCAATGACCTTGTGCT	5347
Db	5837	AGGAGAGATCAAGATGCTGCTCTCTGTGTGTGAGCCCAAGACCTTTATGACCTTCCTTT	5896

QY	5348	CCAGGTAGAGACACCGGCTGGGCAAGGACCTGGCCCTCCAAACAACTGCTGAATCCAGG	5407
Db	5897	CCAATCAGAGSACCCGACTAGGCAAGACTCACCCTTCACACAACTGCTCTATGCCAAG	5956
QY	5408	ACATCCCCAGCTACAGAAATTTGGTGGAGAGGATTTACTCAGACATAGGGAGATGCCAG	5467
Db	5957	ATATCCCCAGTTAATAAGACTGGGTATAAAGATCTATGAGATATTTGCCAAGCTCCAG	6016
QY	5468	CCATCAGGACCAAGACATGACCCATACCTGGCTGACAGTCCCGGATGCATCATGATG	5527
Db	6017	CCATTAGTAGCAAGATATGATATGCTCTACCTCGGAGAGAGTCCGCTGGATGCTACAG	6076
QY	5528	AGTTCAACACCATGATGTCACACTCTTCAGAGATCTTCTCTATGTGGGCAATATACGCGAG	5587
Db	6077	AGTTCAATATGCTAGAGCGCCCTCAACAGATCTTACTATGTGAGCAAGTACGTAGG	6136
QY	5588	AGATTCCTTGACCTCTGGACCCAGCATGACCAAGTGTGGAAAGCAGAAACTGGCTTACAAC	5647
Db	6137	AGCTCATGGGGGCACTAGAGCAGGATGAAACAGGGCCGACCGCAACGACTGGGCTTACAAG	6186
QY	5648	TAGAACAAAGTCATAACCTCATGAGCTTTAGACAGCTGAATATAA	5691
Db	6197	TGGAGCATTCTCATCAAGCCATGCTGCATATGAGACTTAAAGGAA	6240

AX704766	6147 bp	DNA	Linear	PAT 04-APR-2003
RESULT 8				
AX704766				
LOCUS				
DEFINITION	Sequence 70 from Patent WO0202634.			
ACCESSION	AX704766			
VERSION	AX704766.1			
KEYWORDS	GI:29561432			
SOURCE				
ORGANISM	Homo sapiens (human)			
REFERENCE	Homo sapiens			
AUTHORS	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo.			
TITLE	1			
JOURNAL	Gururajan, R., Hafalja, A.J., Kallijk, D.A., Patterson, C., Azimzai, Y., Khan, F.A., Xu, Y., Yao, M.G., Yue, H., Au-Yang, U., Batta, S., Baughn, M.R., Borowsky, B.A., Io, T.P., Lu, D.A., Lu, Y., Tang, Y.T., Yang, U., Zingler, K.A., Deleage, A.N., Glezer, K., Marcos, G.A., Nguyen, D.B., Policky, U.L., Ramkumar, V., Thangavelu, K., Walla, N.K. and Warren, B.A.			
FEATURES	Human extracellular matrix and cell adhesion polypeptides			
source	Patent: WO 0202634-A 70 10-JAN-2002;			
	INCYTE GENOMICS INC. (US)			
	Location/Qualifiers			
	1..6147			
	/organism="Homo sapiens"			
	/mol_type="unassigned DNA"			
	/db_xref="taxon:9606"			
	/note="IncYTE ID No: 7204554CB1"			
ORIGIN				
Query Match	46.7%; Score 2659.4; DB 6; Length 6147;			
Best Local Similarity	68.4%; Pred. No. 0;			
Matches 3811; Conservative	0; Mismatches 1716; Indels 42; Gaps 8;			
141	CGCCGAGGTTTCAATACACCTCGTGGGAGAGAGAGACAGACACAACTTACCTTGGGGGC	200		
302	CGACTGGGGGCTCCACCCACTAGTGGTCATGACACACAGACGACGAGTGTATGTGGGGCC	361		
201	CGTCATTCGATTTCACAGCTCTCCAGGACCTGAAGGTTTGGTGAACGATGACAG	260		
362	AGTGAACCGCATCTATTAAGCTGTGGGGAACTTACACTGCTGGGGCCACGTCAACGG	421		
261	GCCGAGACAGAGACCAACCCAGAGTGTTCACCCAGCCGCGATCGTCCAGACCTGCAATGAGCC	320		
422	CCCTGTGAGAGACAAACGAGAGTGTCTACCCGCGCCAGACGCTGAGTCTCTCCCCACGG	481		
321	CCTGACCAACCAACAAATGTGAACAGATGCTCTCATTAAGCTTCAACAGAGGAACAGGCT	380		

Db	482	CCTGGGCAAGTACTGACAAACGTCAACAACACTGTGTGCTGTGACATATGCGCTAACCGGCT	541
Oy	381	GATTGCTGTGGAGCCTCTGTACCAAGGCAATTCGAAGCTGTGAGGCTGGAAGACCTCTT	440
Db	542	GCTGGCCCTGTGGCAGCGGCTCCCAAGGGCATCTGGCCAGTTCCTGCTGTGAGCAATCTCTT	601
Oy	441	CAAGTGGGGGAGCCTTATCATTAAGAAAGACATATGTCAGGTGTCCATCGAAGCGG	500
Db	602	CAACTGGGTGAGCCACACACCGTAAAGAGCATCTCTGTCCAGCGTGCAGAGGGCAGG	661
Oy	501	CTCACTCTTTGGAGTGATCGTCTCTAC-----AGCAACTGTATGACAAAGCTGTTCAT	554
Db	662	CAGCATGGCGGGCGGTGCTATTGGCGGGCCACCGGGCCAGGGCCAGGCCAAGCTCTTGCT	721
Oy	555	TGCCACGGCAGTGAATGGGAAGCCGAGATATTTTCCACACATCTCCACCCGAAACATGAC	614
Db	722	GGGCACACCCATCGATCGAATTCGAGATCTTCCCAACCTGTCCACACCGCTGGCTCAT	781
Oy	615	CAAGAATCTGAGGAGGATGGCATGTTGCGCGTACGCTTCCATGATGATGAGTTGCTGGGCTC	674
Db	782	GGCCACACAGAGAGATGCGGCAATTTTGCGTTCGTGTATACAGATGATGTTGTGTATC	841
Oy	675	GATGATTAGATCCCTTGGACACTTCAACCATCATCCCTGAATTGATATCTACTATGT	734
Db	842	ACAGCTCAAGATCCCTTCGGACACCGCTGCCAAGTTCGCGCTTTTGACATCTACTATGT	901
Oy	735	CTATGTTTTAGCAATGGCACTTTGTGTACTTTTGAACCCGCCAAGCTGATGATGTGC	794
Db	902	GTAAGCTTCGCGACGAGCACTTGTGTACTACTCAGCGCTGACGCTGACACACAGCT	961
Oy	795	TCCACCAAGCTTCACCCACCAAGAGCAGGTGTATCATCAAGCTGTGAGGCTTTGCAA	854
Db	962	GACCTCGGCTGATGCGCGCGGGACCATTTCTTCAAGTCCAAAGATCGTGGGCTGTGCT	1021
Oy	855	GGAGGACACAGCTTCAACTCCTATGTAGAGTGGCCATTGGCTGTGACCGGAGTGGGT	914
Db	1022	GGACACCCCAATTTCTACTGTATGATGAGTTCCCAATGGCTGTGACGACAGCGGGTGT	1081
Oy	915	GGAGTACCGGCTGCTGAGAGCTGCTCACTCTGTCCAAACCGGGGAGCGGTGTGACAGAC	974
Db	1082	GGAATACCGGCTGTGAGAGATGCTTACTGAGCGGGCGGCGGCTGTGGGCCACACA	1141
Oy	975	CCTTGGATGCATCCAGATGATGACCTGTCTTCAACGCTCTTCTCMAAGGCGAAGACG	1033
Db	1142	GCTGGGCTGTGGTGAAGACGAGAGGTGTCTTCACTGTGTTCGCCACGGGCGCAAGAA	1201
Oy	1035	GAAAAATGAATCCCTGAGTGAAGTCCGACCTGTGCATCTTCAATCTTGAAGACAGATTAATGA	1094
Db	1202	CCGCGTGAAGCCACCAAGAGATCAGCACTGTGCTGTTCACGCTCAGGCGCATATAGGA	1261
Oy	1095	CCGACTTAAAGAGCGGCTGCACTCTTGTAAACCGGGGCGAAGGCAAGCTGAACCTGGCGTG	1155
Db	1262	GAAAGTTAAAGAGCGCATTCAGTCTCTGCTACCGTGTGAAGGCAAGCTCTCCCTCGGTG	1321
Oy	1155	GCTCAAGGTGAAGACATCCCTGACAGATGTGCGCTCTTAACATGACGATTAATTTG	1214
Db	1322	GCTGTCTAAACAAGAGCTGTGGGTGATCACTGCGCCCTGCAGATCGATGACATCTTG	1381
Oy	1215	TGGCTGTGACATGAATGTCTCCCTGAGAGTGTCCACATGTGTGCGTGAATATCCCGTCT	1274
Db	1382	CGGGAGAGACTTCAACAGCGCCCTTGGGGGCAATCAACATTAAGGGGAGCGCCCTGT	1441
Oy	1275	CACGAGAGACAGGAGCCGATTAAGTCTGTCAATGCAATATGTCTACAAAGAACATCTCT	1333
Db	1442	CGTGACAAAGATGTATGCTGTACCGCGGTGCTCCTATGACTATCGGGGCGGACATGT	1501
Oy	1335	GAGCTTTTGGGCAACAAGTGGCAAGGTGAAAGAAAGATCCGGGTGATGACCA-----	1390
Db	1502	GGTATTTCGCGGCAAGCAAGTGGCGCATCCGCAAGATCTGTGTGAACCTTCAAAACC	1561
Oy	1391	-----GGGGCAACGCCCTTCAGATATGACGGGTGACGTTGTGTGACCCCGGCCACGTCT	1441
Db	1562	CGATGCGCGGCTGTGCTGCGCTTACGAGAGGCTGTGTGCGCCAGAGAGGAGCGCCCATCT	1621

[illegible]

QY 2523 GGAGAGCCAGTGTGAGTGTCTGTGTGTCGCAAAAGCAATGCAAAACCCCGCATAC 2582
DB 2702 ACTGTCATCTGTGATGACGCGCGTCAAGGCAAGTCGTCACCCGACCCCAAGTCTT 2761
QY 2583 AGAGATAATCCCGGTGACAGGCCCCCGGGAAGGGGGACCAAGTCACTATCCGAGGGGA 2642
DB 2762 CAAGTGTCTCCCGAGAGCGGCCCCGAGGAGGGGCGCGCGCTCACTATCACAGGCGA 2821
QY 2643 GAACCTGGAGCTTGAATTTCCGCAATGTGCTCCCAATGTCAAGTTGTCTGCGGTGATG 2702
DB 2822 GAACCTGGAGCTTGAATTTCCGCAATGTGCTCCCAATGTGCTCCCAATGTGCTG 2881
QY 2703 CAGCCCTTGAAGTGTGATTCATCCCTGACAGAAACAGTCTGTGTGATGATGGGGAGGC 2762
DB 2882 CAGCCCTTGAAGTGTGATTCATCCCTGACAGAAACAGTCTGTGTGATGATGGGGAGGC 2941
QY 2763 CAAGCCC--AGCCAGCATGAGGCTTGTGGAATCTGCGTGGCTGTGTGTGCGGCTGA 2819
DB 2942 CAGCTCCGTGGGTGCGCATGACGCGCTGTGTGAGGTGTGTGTGCGGAGCTGTCAACCA 3001
QY 2820 ATTGATGAGCCGGTCTCTACAGCTCTATTACTTATGACACTGACTCTCTCAATGTGAA 2879
DB 3002 CTACCGGCTCTGTGACCAAGGCTTCACTTGTGTGACCAACCTTCTACCGTGTGAG 3061
QY 2880 GCCCAGCCGGGGCCCATGTCCGAGAGGACCCCAAGTGCATCAAGGCAACCACTGAA 2939
DB 3062 CCCCTCCGTGGGCTCTGTGAGGGGCACTGATTTGGCATTCAGGAGAAAGCCACTGAA 3121
QY 2940 TGGCGGAGAGCAACGTGTGTGTGTGTGTAAGAGCCCTGTCTTTCACAGGCGATC 2999
DB 3122 CGCAGGAGT 3181
QY 3000 TCCATCTTCAATGTGTCTGCAACACCAATCTCAGATGAGTGTGATGATGAGTGTG 3059
DB 3182 CTCCGTGAGATCCGCTGCTGTGACACCCCGGAGAGCCCTGGCAGCGCTCCCATAT 3241
QY 3060 GGTGAGGTGAGCAGGAGGCAAG--ATCCACAGAGCCTGTCTTCAATGTGTGAA 3116
DB 3242 CATCAACATCAACCGCGCCAGCTCACCAACCTGTGAGTGAATGATCACTACCGAGGA 3301
QY 3117 CCCCACATCTGTGCGAATTGAGCCAGAAATGAGCATTTGTCAGTGAAACACACCCATG 3176
DB 3302 CCCCACATCTGTGCGAATTGAGCCAGAAATGAGCATTTGTCAGTGAAACACACCCATG 3361
QY 3177 CGTATGGGGAGCCCATCTGTGACCTATACAAACCCCAAGTCCGTGCAAGCATGAGG 3236
DB 3362 GGTCAAGGCAACCTGTGCGCATGTCTGAAACCCGAAATCCGGGCAAGTGTGAGG 3421
QY 3237 GAAGAGCAGCATATCTGTGAGGTCTGTGACGCTACTGAGATGACCTGTCAAGGCGC 3296
DB 3422 CATTTAGAGGAGGAACGGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 3481
QY 3297 CGCCCTGTCTGGGTCTGTGACCACTGACCACTGACCACTGACCACTGACCACTGAC 3356
DB 3482 GTCTGTGTGCAACCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 3541
QY 3357 CTTCATCTGTGACAGTGTGAGTCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 3416
DB 3542 CTTCATCTGTGACAGTGTGAGTCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 3601
QY 3417 TCCCAACCGGT 3476
DB 3602 CCTTACCCCGTACTGTGAGGACACTCAGCCCACTGCGCTGTGTGTGTGTGTGTGTGT 3661
QY 3477 GGCATCATCTTAAAGGGGAGAACTGTATCCCGCTGTGTGTGTGTGTGTGTGTGTGT 3536
DB 3662 CCTTACCTTCTCAAGGGGCGGAACTCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 3718
QY 3537 GAATCACTGT 3596
DB 3719 CAATCACTGT 3778
QY 3597 GGTGTGAGT 3656

DB 3779 GGTGTGAGT 3838
QY 3657 GAGATATCTCCCGGGGATGTGTATCTTGTCCCGGACAGCCCGCTCACTGTCCCGCAT 3716
DB 3839 CGAGTCTTGTGCGAGGGAACCTGTGAGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 3898
QY 3717 CGTGTGATCTGT 3776
DB 3899 TGTGTGATGT 3958
QY 3777 CTATTAAGCAAGTCTCCGGAATGTACCTCACTGTGAAGCGGTGTGAGATGTGATGGA 3836
DB 3959 CTACAAAGCGCAAGTCAAGATGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 4018
QY 3837 CAACCTGTGAGT 3896
DB 4019 CAACCTGTGAGT 4078
QY 3897 CATTCATGAGT 3956
DB 4079 CATTCATGAGT 4138
QY 3957 TTACACATGT 4016
DB 4139 ATATGT 4198
QY 4017 GGT 4076
DB 4199 GGT-----GAGGCGCAATGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 4246
QY 4077 CAACCAAGAGT 4136
DB 4247 GACCAAGAGCACTTGT 4306
QY 4137 CATGTGAGCAGT 4196
DB 4307 CATGTGAGCAGT 4366
QY 4197 GTATGCGCATGT 4256
DB 4367 ATATGCGCATGT 4426
QY 4257 CAAGAACGACCTTAAGT 4316
DB 4427 CAAGAACGACCTTAAGT 4486
QY 4317 CAATTGT 4376
DB 4487 TAATGT 4546
QY 4377 CTCCGT 4436
DB 4547 CATGT 4606
QY 4437 CGAGGCGCGCTACTCTGT 4496
DB 4607 TGAAGGACGCTACTCTGT 4666
QY 4497 CTTGT 4556
DB 4667 ACTGACCTTGAATGT 4726
QY 4557 CTTCAACTGT 4616
DB 4727 GCTGT 4786
QY 4617 TGT 4676
DB 4787 CGT 4846
QY 4677 TGGGCGAAGATGT 4736

Db 4847 CATGGCGGCATCATCTGTCAGAGACGAGACGTCACCAAGATTGACCAAGATTGAA 4906
 QY 4737 GGGACGGAACACACCTGGCCCACTACCAAGGTGCGAAGTGGTCCGGTGGCATGATGTC 4796
 Db 4907 GAGGCTGAACACACCTGGCTCACTACCAAGGTGACAGAGGCTCTGGTGGCAGCTGTC 4966
 QY 4797 CAAGCAGGTGACAGCTTATACGAGTGAACAATCCACGCTCTCCAGAGCTTGACAG 4856
 Db 4967 CAAGCAGAGCTGCGCTTACACATCTCCAACTCTCCACCTTACCAAG---TCCCTCAG 5023
 QY 4857 TAAATATGAATAATATATCCGATACCGGCGAGCCCCGACGCTCCGCTCAGGACACC 4916
 Db 5024 CAATATCGAGAGATCTGTGCGACGGCCAGCAGCCCCGACGCTGCGCTCAGGACGCC 5083
 QY 4917 TATGATCACTCTGCTGAGAGTGAAGTCAAGATGTCGACCTAGTGAAGACACAGA 4976
 Db 5084 CATGATCAAGCCCGACCTGGAGAGGGGACCAAGCTGTGCACTGGTGAAGACACAGA 5143
 QY 4977 GCAACGAGACCAAGAGAGGGGACCCGGGGAGCAAGATGTGTCTGAATCTTACTGAC 5036
 Db 5144 CCACTGAGACCAAGCTGAGGGTGAACGCGAGCAAGATGTGTCTGAGATCTTACTGAC 5203
 QY 5037 CCACCTCTGAGCACTAAGGGGCACTGAGAGATTTGTGATGACCTTTTGAAGACAT 5096
 Db 5204 AAGGCTACTGGCCACCAAGGGGCACTGAGAGATTTGTGAGAGACTGTTTGAAGACAT 5263
 QY 5097 CTTGACGAGGACACCGTGGCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 5156
 Db 5264 CTTGACGAGGACACCGGAGGCTGACGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 5323
 QY 5157 CTTGAGTGAAGAGGCTGATTAACATGATGATGATGATGATGATGATGATGATGATGAT 5216
 Db 5324 CTTGAGTGAAGAGGCTGATTAACATGATGATGATGATGATGATGATGATGATGATG 5383
 QY 5217 GAGCAATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 5276
 Db 5384 GAGCAATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 5443
 QY 5277 TGAATCACTTAAGAGAGATGATGATGATGATGATGATGATGATGATGATGATGAT 5336
 Db 5444 CGACATTCACAGAGAGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 5503
 QY 5337 GGAATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 5396
 Db 5504 GGAATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 5563
 QY 5397 GTATGCGAAGAGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 5456
 Db 5564 GTATGCGAAGAGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 5623
 QY 5457 GAAGATGCGAGGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 5516
 Db 5624 GAAGATGCGAGGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 5683
 QY 5517 GCAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 5576
 Db 5684 GCAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 5743
 QY 5577 ATACAGGAGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 5636
 Db 5744 GTACAGGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 5803
 QY 5637 GGCCTACAACTGAGAACTGATTAACCTGATGATGATGATGATGATGATGATGAT 5685
 Db 5804 GGCCTACAACTGAGAACTGATTAACCTGATGATGATGATGATGATGATGATGAT 5852

RESULT 9
 LOCUS D86948 6466 bp mRNA linear ROD 31-MAR-2000
 DEFINITION Mus musculus mRNA for plexin 1, complete cds.
 ACCESSION D86948
 VERSION D86948.1 GI:1665756

KEYWORDS SOURCE ORGANISM

plexin 1.
 Mus musculus (house mouse)
 Mus musculus
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

REFERENCE

1 (sites)
 Kameyama, T., Murakami, Y., Suto, F., Kawakami, A., Takagi, S.,
 Hirata, T., and Fujisawa, H.
 Identification of a neuronal cell surface molecule, plexin, in mice
 J. Cell Biol. 148 (6), 1283-1293 (2000)

TITLE

Determination of cell adhesion sites of neuropilin-1
 J. Cell Biol. 148 (6), 1283-1293 (2000)

JOURNAL

20191894
 MEDLINE
 PUBMED
 10725340

REFERENCE

3 (bases 1 to 6466)
 Fujisawa, H.
 Direct Submission
 Submitted (07-AUG-1996) Hajime Fujisawa, Nagoya University Graduate
 School of Science, Division of Biological Science, Furo-cho,
 Chikusa-ku, Nagoya, Aichi 464-01, Japan
 (E-mail: fujisawa@bio.nagoya-u.ac.jp, Tel:052-789-2978,
 Fax:052-789-2979)

FEATURES

Location/Qualifiers
 1..6466
 /organism="Mus musculus"
 /mol_type="mRNA"
 /db_xref="taxon:10090"
 /tissue_type="brain"
 /dev_stage="E17.5"
 265..5948
 /codon_start=1
 /product="plexin 1"
 /protein_id="BAA13188.1"
 /translation="MPLPLSRLTLLLLLLLRGVIAISSPAGGPGAPARTFVA
 SDMGITLVVHGTGQVAVGAVNRIYKLSGNTLLRAHTGPDENKCYPPSVSC
 PHGSGTDVNVKLLIDYANRLIAGSASQGIQCTRLDPLFKLEPHRRKHYSS
 VREGSGAVGLINGPQGGAKLVGPIFGKSEFTPLSSRLYANEDADMGVY
 QDEFVSQNLPSIDTUSKEPAFDIYVYSRSQFYVYLLQDPLDTSDAGERF
 TSXIVRLQVNDPKFYSEYEPICQEGYEVRLVQAYLSRPGQALAKDGLADEV
 LFVFAQGNRYKPKESALCFTLRALKEIKERIOSCYRGEGLSLFWLNLKEIG
 CINSPLQIDDFGQDNPOLGTVTEGTFLEVDKEDGLTAAVADYQGRVYFAGT
 RSGRIKTIYDNLPSGRPALAYESVVAQGNILDDIVSPRYVAMTEQVQV
 PVESCVQYTCSELCISRPDHCQVYLSICSDQDCERAPRSPASDLLQCVLT
 QPRSVYVMSQVPLVLOANVPDLISAGVNSCFEDFTBESILDGRIHCSPEAREYA
 PITQGGDDQVFLVLYKSKETGRKFAVDVFNCSVHOSCLACVNSPFCBCKRKH
 VCTNNADCAVFLGRVMSQDQPIILPSTHIVYVAVVKEITLAAANLPDQSGCKRY
 NLOAHLFGSPARVYALRFNSSLQCONS SVSYSGNVSQDLFWLVYVNGNPFYIDNPQ
 NIOAHLKCPALROSGLCLKADPREFCGVNERCSLHHPCDPSASMMAHHGS
 SRGDDPKLLKSPETGPRGGTRLITIGNGLGRFEDVRLGVAVGKVLCSPVSEYIS
 ABOIVCEIGDASTLRADALAYECVADCSLHYRALSFKRPTFTVTPHYRVSRSGLS
 GGTWIGTBSHLNAGSDVAVSISGRPOSREMRNSREIRCTTPRGHTPGSPFYINIR
 AQLSNPEKXNYVEDPTILRIDPWSINSQTLVYVGNLATVREPRIRAKGGIER
 ENSCMVYVNDTVMVCRAPSIDNPKRSPPELBERDEIGFINDVNTLLVNSSFLYP
 DPVLEPLSPGLLEKRSPLILKGNLIPAGNSLANTVYIGSPCLLVSEVQL
 LCEAPNLTPGQKTVRAGGEPSPGMQVYSQDLTLPAVIGGGGGLLVAVL
 IAYRKGRADRLTKLOMDNLESVALLCEKFAPELOTDIHELTLSDAGIPL
 DRYTYRNYVPRGIEDHPUTKMEVQANVKSLLTGQQLTKHGFIILTPTRLEAOS
 FMSDRGNVSLIMTALQGMETATGLKLSLDLLEKLSGNSRHPRLIRRESVAB
 KMTLNTFTPLVFLKEAGEPFLVYCAIKQOMEGSPDITAGEARYSLSERLHQ
 QIDVKTLLINCVNPEHENABEVKGLNCDIVYQVEKLDADAYKVPISQREKAGM
 DLEWQGRMARILLQDEDTYTKIDNMRKNTLADYQVDSVVALVPKQTSYVNS
 SSTFTKLSRSESLRTASPSDLSRSTPMITLDSGTGLMLVKNHMLDPRBGR
 GSKVSEIYVTRILATKGLQKQVDFETFTSAHGSALPLAIKMPFLBOADK
 HOIHDSDVRETKWNSCLPLRFVNYNIPQFVDIKHNSITDCLSVATQFMDSCST
 SEHKLQDQSPSKLIYAKDIPYKSKVTERYVADIAPKAPASIDQMSVYLAEGSRHLS
 QFNSMSALHETIYIAIKDEILVALLEKQDARRRRLKLEOVVDMALS"

FEATURES

Location/Qualifiers
 1..6466
 /organism="Mus musculus"
 /mol_type="mRNA"
 /db_xref="taxon:10090"
 /tissue_type="brain"
 /dev_stage="E17.5"
 265..5948
 /codon_start=1
 /product="plexin 1"
 /protein_id="BAA13188.1"
 /translation="MPLPLSRLTLLLLLLLRGVIAISSPAGGPGAPARTFVA
 SDMGITLVVHGTGQVAVGAVNRIYKLSGNTLLRAHTGPDENKCYPPSVSC
 PHGSGTDVNVKLLIDYANRLIAGSASQGIQCTRLDPLFKLEPHRRKHYSS
 VREGSGAVGLINGPQGGAKLVGPIFGKSEFTPLSSRLYANEDADMGVY
 QDEFVSQNLPSIDTUSKEPAFDIYVYSRSQFYVYLLQDPLDTSDAGERF
 TSXIVRLQVNDPKFYSEYEPICQEGYEVRLVQAYLSRPGQALAKDGLADEV
 LFVFAQGNRYKPKESALCFTLRALKEIKERIOSCYRGEGLSLFWLNLKEIG
 CINSPLQIDDFGQDNPOLGTVTEGTFLEVDKEDGLTAAVADYQGRVYFAGT
 RSGRIKTIYDNLPSGRPALAYESVVAQGNILDDIVSPRYVAMTEQVQV
 PVESCVQYTCSELCISRPDHCQVYLSICSDQDCERAPRSPASDLLQCVLT
 QPRSVYVMSQVPLVLOANVPDLISAGVNSCFEDFTBESILDGRIHCSPEAREYA
 PITQGGDDQVFLVLYKSKETGRKFAVDVFNCSVHOSCLACVNSPFCBCKRKH
 VCTNNADCAVFLGRVMSQDQPIILPSTHIVYVAVVKEITLAAANLPDQSGCKRY
 NLOAHLFGSPARVYALRFNSSLQCONS SVSYSGNVSQDLFWLVYVNGNPFYIDNPQ
 NIOAHLKCPALROSGLCLKADPREFCGVNERCSLHHPCDPSASMMAHHGS
 SRGDDPKLLKSPETGPRGGTRLITIGNGLGRFEDVRLGVAVGKVLCSPVSEYIS
 ABOIVCEIGDASTLRADALAYECVADCSLHYRALSFKRPTFTVTPHYRVSRSGLS
 GGTWIGTBSHLNAGSDVAVSISGRPOSREMRNSREIRCTTPRGHTPGSPFYINIR
 AQLSNPEKXNYVEDPTILRIDPWSINSQTLVYVGNLATVREPRIRAKGGIER
 ENSCMVYVNDTVMVCRAPSIDNPKRSPPELBERDEIGFINDVNTLLVNSSFLYP
 DPVLEPLSPGLLEKRSPLILKGNLIPAGNSLANTVYIGSPCLLVSEVQL
 LCEAPNLTPGQKTVRAGGEPSPGMQVYSQDLTLPAVIGGGGGLLVAVL
 IAYRKGRADRLTKLOMDNLESVALLCEKFAPELOTDIHELTLSDAGIPL
 DRYTYRNYVPRGIEDHPUTKMEVQANVKSLLTGQQLTKHGFIILTPTRLEAOS
 FMSDRGNVSLIMTALQGMETATGLKLSLDLLEKLSGNSRHPRLIRRESVAB
 KMTLNTFTPLVFLKEAGEPFLVYCAIKQOMEGSPDITAGEARYSLSERLHQ
 QIDVKTLLINCVNPEHENABEVKGLNCDIVYQVEKLDADAYKVPISQREKAGM
 DLEWQGRMARILLQDEDTYTKIDNMRKNTLADYQVDSVVALVPKQTSYVNS
 SSTFTKLSRSESLRTASPSDLSRSTPMITLDSGTGLMLVKNHMLDPRBGR
 GSKVSEIYVTRILATKGLQKQVDFETFTSAHGSALPLAIKMPFLBOADK
 HOIHDSDVRETKWNSCLPLRFVNYNIPQFVDIKHNSITDCLSVATQFMDSCST
 SEHKLQDQSPSKLIYAKDIPYKSKVTERYVADIAPKAPASIDQMSVYLAEGSRHLS
 QFNSMSALHETIYIAIKDEILVALLEKQDARRRRLKLEOVVDMALS"

ORIGIN

D86948 6466 bp mRNA linear ROD 31-MAR-2000
 Mus musculus mRNA for plexin 1, complete cds.
 D86948.1 GI:1665756

Query Match 45.3%; Score 2578.4; DB 10; Length 6466;
Best Local Similarity 67.5%; Pred. No. 0;
Matches 3760; Conservative 0; Mismatches 1766; Indels 42; Gaps 8;

QY 142 GCCGAGGTTTCAATCACTGTGTGTGATGAGAGACAGACATTTACTTGGGAGCC 201
DB GACTGGGAGCTCACTCACTGGTGTTCACGAAACAGACAGAGATATATGTGGCCT 459
QY 202 GTCAATCGATTTTACAGCTCTCCAGCCAGCTGAAAGTCTTGTATACGATAGACAGG 261
DB 460 GTGAACCGTATTTTACAGCTGTGGGAACTTAACTGCTTGGGAGCTATGTGACAGC 519
QY 262 CCGGAGAGGACAAACCCAGTGTTCACCAACCCGATGCTGACAGCTGCAATAGGCC 321
DB 520 CCGTGAAGACATAGAAATGTCTACCAACCCCGAGTGTACAGTCCGACCAATGT 579
QY 322 CTGACCAACCAACATGTCAACAAAGATGCTCTCATAGACTTACAGAGAAACAGGCTG 381
DB 580 CTGGGACAGACAGACAGTCAACAGCTTCTTACTGTGACTATGACGCCAACGCGCTG 639
QY 382 ATTGCTGTGGAGCTGTACCAAGCATGTGCAAGCTGCTGAGGCTGAGGACCTTTC 441
DB 640 TTGGGCTGTGGAGTGTCTCAAGGCACTGTGCACTTCTGAGGTGGACATCTCTTC 699
QY 442 AAGCTGGGAGGCTTATATTAAGAGACATATCTGTCAAGTGTCAACGAGAGCGGC 501
DB 700 AAGCTGGGAGGCGGACACACCGCAAGGAACTACTGTCCAGTGTCAAGAGGCTGGC 759
QY 502 TCACTTGTGAGTATGCTCTCTA-----CAGCAACCTGGATGACAAAGCTGTCTT 555
DB 760 AGATGCGGAGATGTCTATGTGTGGCCACCCGGCAGGGCAGGCAAGCTCTTCTGT 819
QY 556 GCCACGCGAGTGTAGTGGAGGCCCGAGTATTTTCCACCATCTCCAGCCGAGAACTGACC 615
DB 820 GGGACACCCATGATGTGGCAATACAGAACTCTCCCACTGTCCAGTCCGCGCTCATG 879
QY 616 AAGAACTGTGAGCGGATGAGCATGTTGCGTACGCTTCCATGATGAGTGTGGCGG 675
DB 880 GCCACAGAGAAATGTCTACATGTTTGGCTTGTGTACAGAGATGATGTGTCTCATCT 939
QY 676 ATGATTAAGATCCCTTGGACACTTCAACATCATCTGACTTGTATATCTATATGTC 735
DB 940 CAGCTCAAGATCCCTTCAACACATTTGTCAAGTTCACGAGCTTGTGACATCTATATGTG 999
QY 736 TATGTTTACAGATGAGCACTTGTCTCTTTTGAACCTTCCACCTGAGATGTGTCT 795
DB 1000 TACAGTTTCGAGCGAGCAAGTGTGTATCTACTCACTGCACTGACCTAGACACGCAAGCTT 1059
QY 796 CCACAGGCTCCACACCAAGAGAGAGTGTATACATCCAAAGCTGTGAGGCTTTGCAAG 855
DB 1060 AACTGCCCGGAGCGGCTGGGAGCACTTCACTCCCAAGATGTGGGCTGTGGCTG 1119
QY 856 GAGAGACAGAGCTTCACTCTCTATGTAGAGTGGCCATTTGGTGTGAGCGCAGTGGAGTG 915
DB 1120 AATGACCCCAAGTCTTACTTCTTACGTGAGTTCCTCCATTTGGTGTGAGCAGCTGGCGTG 1179
QY 916 GAGTACCGCTGTGAGAGCTGCTCACTGTCCAAAGCGGAGCGCTGTGGCAGAGCC 975
DB 1180 GAGTATCGCTGTGTGAGAGCGCTTACCTGAGCCGCGGAGCAGAGCCCTGGCCAAAG 1239
QY 976 CTTGAAGTCAATCAAGATGATGACTGCTCTTCAACCGTCTTCCAAAGGCGCAAGAGCGG 1035
DB 1240 CTGGGCTGTGTGAGATGAGAGGTGTGTTCATGTGTTCGACAGAGGCGCAAGAAAGAC 1299
QY 1036 AAAATGAATCCCTGGATGAGTGGGCTGTGATCTTCACTTGTGAAGAGATTAATAC 1095
DB 1300 CCGGTGAAGCCACCAAGAGAGTGGGCTGTGCTGCTGTTCACACTGAGGCCATCAAGAG 1359
QY 1096 CGCATTAAGAGAGCGCTGCAAGTCTTGTTCACCGGAGCGAGGCAAGCTGTGAGCTGTGG 1155
DB 1360 AAAATCAAGAGAGGAGCTCAATCTGTGTACCGTGAAGAGAAAGAACTCTCTGCTGTGG 1419
QY 1156 CTCAAGGTGAAGACATCCCTGACAGAGTGTGCTTCAACCATTTGACGATTAATCTTGT 1215

DB 1420 CTGCTCAACAAAGAGTGGCTGTGATCAATCCGCCCTGTGAGATTTGATGACATCTTTC 1479
QY 1216 GGCCTGACATGAATGCTCCCTGTGAGAGTGTCCGACATGCTGGTGGAAATTCGCTTC 1275
DB 1480 GGGCAGGATTTCAACAGCCACTGGGTGGACCGGTCAACCATGAGGGCACACCCCTTTT 1539
QY 1276 ACGGAGGACAGGAGCCGCAATGAGTGTGTATGCAATGTCTTCAAGAAACCATCTCTG 1335
DB 1540 GTGACAGAGAAATGTGCTGACTGTGTGCTGTGCTGTCTATGATCAACAAAGGCGCACCTG 1599
QY 1336 GCTTTTGTGGCAACAAAGTGGCAAGCTGAPAGAAATCCGGTGTGATGACCCAGAGGC 1395
DB 1600 GTTTTGTGCACTGTGACAGCGCCGATCCGCAAGTTCTGTGTGACCTTGCACAAACCC 1659
QY 1396 AAC-----GCCCTCCAGTATGAGACGCTCAGTGTGTGACCCCGGCCAGTCTTC 1446
DB 1660 AGCGGAGCGCGCCCTGTGTATGAAGTGTGTGTGTGAGAGGAGCAACCTATCTGTG 1719
QY 1447 CCGGATATGAGCTTCTTCAAGACACAGCACTCTATCATGTCTAGAGAGCAGCTC 1506
DB 1720 CGTGACCTGTGCTTCAAGCCCAACCGCAGTACTCTTATGTATGAAGAGAGAGGTG 1779
QY 1507 ACCAGATCCCTGTGAGTCTGTGTGCTGATCAAGAGCTGCGGCAAGTGTGCTTGTGCTCA 1566
DB 1780 ACCGAGGTGCTGTGAAAGCTGTGTCCAGTATATCTTGTGAGCTGTGCGGAGTCT 1839
QY 1567 GCGACCCCACTGTGAGTGTGTGTGTGTGCAACACGTGACCCGGAAGAGCGAGTGT 1626
DB 1840 CCGGAGCCCACTGT 1899
QY 1627 GAGCGGTCAAGAGAGCCGAGAGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1686
DB 1900 GAGCGGCAAGAGAGCCCAAGCTTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1959
QY 1687 GTCCATCCCAACATATCTCCGTCTCAAGTACAGTGTGTGTGTGTGTGTGTGTGTGTGT 1746
DB 1960 GTCCAGCCGCGCAATGT 2019
QY 1747 AATGTCCGAGAGT 1806
DB 2020 AATGTGTCTACCTGT 2079
QY 1807 GGGCTGT 1866
DB 2080 AGCATCTGTGAAGATGTGCGCATGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 2139
QY 1867 ATCATCAAGAGATGGGAGCCACCATGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1926
DB 2140 ATCAACAGAGGTGAGGAGACAGCGGAGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 2199
QY 1927 GGGATGACCTTGGCAGACAGAGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1986
DB 2200 GGGAGGAGT 2259
QY 1987 CTGTCTGT 2046
DB 2260 CTGGATGT 2319
QY 2047 CATGACCCCAAGACTGT 2106
DB 2320 AACATGT 2379
QY 2107 CAGCTGT 2166
DB 2380 CAGATCTGT 2439
QY 2167 AAGGCCAAGAACTCCCGCAGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 2226
DB 2440 GTTGCACGGAAGCTGT 2499
QY 2227 ATTCAAGGAGAGAGAGAGT 2286

Dp	2500	ATCCAGGACAGCCGGGCCCGGAGTCACTGCACTGGCGCTTCAACAGCTCCAGCTGAGTGC	2555
Qy	2287	CAGAACACCTTATTCTTATGAAAGGATGAGATCAACAACTGCCGTGAGATTGACA	2348
Dp	2560	CAGAACTCCTGTATCTCTATGAGGGGAATGATGTACAGCAATCTGCCTGTAACTGTGC	2619
Qy	2347	GTCGTGGAATGAGGCACTTCAACATTTGACAAACCCAGCTCGAGATPAAAGTTTCACTCTAC	2406
Dp	2620	GTTGTGTGAATGAGCACTTTGTCTATTGACACCCCGAAGACATCCAGGCCCACTCTAC	2679
Qy	2407	AAGTGTGAGCCATGCTGTGAGAGCTGCGGCGTGTGCTCAAGGCTGACCCAGACTTGCA	2466
Dp	2680	AAGTCCCAAGCTTTGCGCCAGAGTTGTGAGCTCTCTCTCAAGGCGAGACCCACGCTTTGAG	2739
Qy	2487	TGTGCTGTGTGCAAGGCGCCAGGCGCAAGTGCACCTCGGCGCGAGACCTGGC---CTGCCAG	2523
Dp	2740	TGCGCTGTGTGTGATGTAAGGCGCGCTGTTCCTTCTGTACCACTGGCCAGCTGATCA	2799
Qy	2524	GAGAACCACTGGCTGAGACTGTCTGTGTCCAAAAGCAAGTGCACAAACCCCGCATCA	2583
Dp	2800	CCAGCTCCTGGATGCATGCCACCATGCGAGAGCGCGCTGACAGACCCCAAGTTCTC	2859
Qy	2584	GAGATTAATCCCGGTACAGGCGCCCGGGAAAGGGGCACCAAGTCACTATCCGAGGGAG	2643
Dp	2860	AAGCTGTCCCAAGAACAGCGCTCTGGAGGAGGAGTCCAGATCAACCATACCGGTAG	2919
Qy	2644	AACCTGGGCTGGAATTTTCGCACTGCCCTCCCATGTCAAGGTTGTGCGGTGAGAGC	2703
Dp	2920	AATCTGGGCTCCGGTTTGAGAGATGTGCGGCTGAGCGGTGCACGTGCGCAAGTCTTTGC	2979
Qy	2704	AGCCCTTATGTGATGTTTACATCCCTGCAGAACAGATGTGTGTAGATGGGGAGGC	2762
Dp	2980	AGCCCTGTGAGAGGAGATATCATCGCCAGAACAGATTGTCTGGAAATTTGAGATGCG	3039
Qy	2763	--CAAGCCACACAGCATGAGGCGCTGTGGAGATCTGCGTGTGTGCGCTGAA	2820
Dp	3040	AGCAGCTGTGCGGCGCCATGATATGCTTGTGTGAGGTGTGTGCGGAACTGCTCTTCA	3099
Qy	2821	TTCAATGCGCCGATCTCTCAAGCTCTATTACTTATGACATCACTGACTCTTCTCATGTGAG	2880
Dp	3100	TACGAGCGCTGTCCCTTAAAGCGTTCACTTTTGTGACACCGACCTTCTACGTGTAGC	3159
Qy	2881	CCCAAGCCGGGGGCCCATATGCTCGGAGAGGACCCAAAGTACATTCAGAGCAACCACTGAAT	2940
Dp	3160	CCCTCCCTGTGGCCCTGTCTGTGAAGTACTGTGAATGTGCATCGAAGGAGCCATCTGAAT	3219
Qy	2941	GCCGGAACCAAGTGTGTGTATGTTTGAAGAGCGCCTGTCTTCTTCAACAGGAGATCT	3000
Dp	3220	GCTGGCATGATGTATGCTGTGTCTGTCCATTTGGCGGTGCGCCCTCTCTTCTCTGAGAAAC	3279
Qy	3001	CCATCTTACATTTGTCTGCAACCAACATCTCTAGTGAAGTGTGCTGAGATGAAGGTGTG	3060
Dp	3280	TCAGAGAGATTCGCTGTGCTGACACCGCCCGGCGCACTCTGTGACGCGCCCTATTGTT	3339
Qy	3061	GTGCAAGTGAACAGGCGCAAGT---CCACCAAGGACCTGTCTTCACTATGTGAAGAC	3117
Dp	3340	ATCAACATCAATCTGTGCCAGCTCTCCAAACCCCGAGTGAAGTACATTAATCTGAAGAC	3399
Qy	3118	CCCAACCATGTGCGGATTTGAGCCAGATGAGACATTTGCTAGTGAACACACCACTGCG	3177
Dp	3400	CTTACTATCTGAGGATGCACTCTGAGTGAAGCATCATTAATGTGGGAGCCCTCTTAACA	3459
Qy	3178	GTAATGGGGAGCCCACTGGAACCTCAATAGAACCCCGAGTCCGTGCGCAACATGAGAGG	3237
Dp	3460	GTACGGGACCAACTGTGCGCATGTGTGTAACCTGTAATTCGGGCGCAATTAAGAGGC	3519
Qy	3238	AAGAGCAATATCTGTGAGGTTTGTGAACGTACTGTGATGACCTGTCAAGCGCC	3297
Dp	3520	ATTGAGAGGAAATAGCTGTGATGTATGATGACCAACATGTGTGTGCGGCGCGC	3579
Qy	3288	GCCCTCGCTTGGGTCTGACCAACAGTCAAGCTGACCGAGAGGCCCGAGAGATTGTGC	3357
Dp	3580	TCATATAGCAACCCAAAGCGAGCGCCGCAAACTGGGAGAGCGCGCAAGATGAATTGTG	3639

QY	3358	TTCAATCCGGAACAAGCCACAGTCCCTGCTCATCTCTCAACAAGACCAATCTTCACTATAT	3417
Db	3640	TTCAATCATGACAATATGACGACAGCTGCTGTACTCAACAGCTCTCTTCTCTACTAC	3639
QY	3418	CCCAACCCGGTGTGTAAGCCCTTGGTCCCTCAGAAATCTGAGCTCAAGCTTGGACG	3477
Db	3700	CCGGAACCTCTATACGTGAGGACCACTTAGCCCTACTGGGCTCTTGGAGCTGAAGCCAGCTCC	3759
QY	3478	CCCATCATCTCAAAAGGGGCAAGACCTGATCCCGCGCTGAGGCTGGGGCAACGTGAACCTG	3537
Db	3760	CCGCTCATCTCTCAAGGGGGCGGAATCTCTGTCACT---GCTCCTGGCAATTCGGGCTC	3816
QY	3538	AATCAACTGTGCTGGTTGGGGAGAGACCGTGACCGTGACCGTGTCAATGTCCAGCTG	3597
Db	3817	AACTAACCGGTGCTCATACGCTCCACACCTTGTATCTCCACCGGTGTGAGACACACACTG	3876
QY	3598	CTCTGCGAGTCCCCCAACCTCATCGGAGGACAAATGATAGGCCCTGTCCGGGGATG	3657
Db	3877	CTGTGTGAAGACCCCAACCTTCACGGGGGACACAAATCAAGGTGCGGGCGCGGCTTC	3936
QY	3658	GAGTACTCCCCGGGGATGTGTATCACTTGGCCCCGAGACGCCCGCTCAGCTCCCGCATC	3717
Db	3937	GAGTCTTCACCTGGGAATGTCTGACAGTGTATTTGGACACGCTGTGACACATCCGGCAAT	3996
QY	3718	GTCAGCAATCGAGTGGGCTGGCGGCGCTCTCATATTTCAATCGGCGCGTCTATTGCC	3777
Db	3997	GTTGGGTATCGGTGGTGGGGGTGGGCTCTTGTCTGTGTATCATGTGGCTGTCTATGCC	4056
QY	3778	TATTAAGCAAGTCCCGGAAAGTGAACCTCACTGCTGAAGCGGCTGAGATGACATGAC	3837
Db	4057	TACAAACGCAATCCCGGAAATGTGACCGTACCTCAAGGGCGTGCAGCTCCCAATGAC	4116
QY	3838	AACCTGGAGTCCCGTGGCCCTTGGAGTGAAGAAACCTTGGCGAGCTGACAGACGAC	3897
Db	4117	AACCTGGAGTACGGGTGGCCCTCGATGTGACGGAACCTTTGAGAGCTACAGACGAC	4176
QY	3898	ATCCATGAGCTGACCACTGATGACTGTGATGAGACCGGGATTCCTGCTCTGACTATAGACT	3957
Db	4177	ATCATATGCTGACCAATGACCTGGAATGGGTGGTGGCACTCCCTCTTGTACTACCGACA	4236
QY	3958	TACACCAATGCGGGTGTCTCCAGAAATTGAAGACCAACCTGTCTCCCGGACCTTGAG	4017
Db	4237	TATGCCATGCGAGTGTCTGTTTCCGTGGATTTGAGGACCATCTGTACTGAAGAGATGGAG	4296
QY	4018	GTCCCGGGCTACCGGCAAGACGTGTGGAAGAGCTGAACTCTTCCGCCACTATC	4077
Db	4297	GT-----GAGGGCCAAAGTGAAGAGTGGCTGACACTGTTCGACAGCTGTGG	4344
QY	4078	AACAACAAGGTGTTCCGTGCTCTCTTATCCGACGCTTGAATGCCAGCTTAGCTTCC	4137
Db	4345	ACCAAGAAAGATTTCTGTTTAATCTTATCCGACACTAAGGCCGACGTAAGCTTCTCC	4404
QY	4138	ATGGCGCAACGTGGCAACGTGGCTCACTATCATATACGTGTGTCAGAGCAAGCTGGAG	4197
Db	4405	ATGTGTACCGTGGAAAGTGGCTCACTATCATGACACCCCTGCAGGGTGTGATGGAG	4464
QY	4198	TACGCCACTGATGTGTGAGACAGAGCTGTGGCGCACTCATTTGACAAGAACTTGAAGC	4257
Db	4465	TATGCCAACAAGAGTGTAAAGCAAGTTGTGTGTGACCTCATTCAGAGAAAACCTTAGAAGC	4524
QY	4258	AAGAAACCACTTAAGCTGTCTCAGAGAGACGTGACTCAATTGGCTGAGAGATGCTGAC	4317
Db	4525	AAGAACCAACCCCAAGCTCTGTTTACGAGAGACGAATCAATGTGGAGAGAAATGTTGACC	4584
QY	4318	AATTGTATTACTTCTCTCTCAATTTCTCTCAAGAGAGTGTGTGGGAGCCCTCTTC	4377
Db	4585	AATGTGTACCTTCTTTGTATTAATTTCTCTCAAGAGAGTGTGTGGGAGCCGCTCTTT	4644
QY	4378	TTCTCTGTTTGTGCAATCAAGACAGACAGATGGAAGAGGCCCAATTGACGCAATCACGGGC	4437
Db	4645	ATGCTGTACTGTGCAATCAAGACAGCAAGATGGAAGAGGCCCTTATGACGCAATCACAGGG	4704

QY 4438 GAGGCGGCTACTCTCTTGAAGGAGACAGCTCATCCGCGCAGATTTGACTACAAACC 4497
 Db 4705 GAGGCGGCTACTCTCTTGAAGGAGACAGCTCATCCGCGCAGATTTGACTACAAACC 4764
 QY 4498 CTGGTCTGAGCTGTGTGAGCCGACAAATGCCAAGCCCGGAGTCCAGTAAATGC 4557
 Db 4765 CTGACCCCTGAAGTGTGTGAGCCGACAAATGCCAAGTCCCGGAGTCCAGTAAAGGG 4824
 QY 4558 CTCAACTGTGACCATCTCACTGAGTCAAGGAGAAATTTGTGATGCCATTTAAAGAT 4617
 Db 4825 CTGAAGTGTGACCATCTCACTGAGTCAAGGAGAAATTTGTGATGCCATTTAAAGAT 4884
 QY 4618 GTGCTGTGCTCCACCGGCGCCAAAGCTCAGATATGATCTGAGTGGCGCAAGAGAGT 4677
 Db 4885 GTGCTGTGCTCCACCGGCGCCAAAGCTCAGATATGATCTGAGTGGCGCGAGGGTCCG 4944
 QY 4678 GGGGCAAGATGATTTTGCAGAGTGAAGACATCACCCAAAGATTGAATGATTGGAG 4737
 Db 4945 ATGGCAGCATCATCTCTTCAAGGAGAGAGCTCACACTAAATTGAATGACTGGAG 5004
 QY 4738 CGACTGAACACACTGGCCCACTACACAGGTGCGAGTGGTTCGATGGGCTTTAGTGTCC 4797
 Db 5005 AGGCTGAACACACTGGCTCACTATCAGTGAAGTGAAGTGTGATGGTGGCACTAGTACC 5064
 QY 4798 AAGCAGGTGACAGCTTATTAAGCAGTGAACAATTCACCGTCTCCAGGACCTGACAGT 4857
 Db 5065 AAGCAGCATCGGCTTACAAAC---ATCTCAACTCTCTTACCTTACATTAAGTCCCTCAGC 5121
 QY 4858 AAATGTGAAACATATATCCGCTGACAGGCGAGCCCGGACGCTCCGCTCCAGGACACT 4917
 Db 5122 AGATTCGAGAGATCTCTGCTGACAGCCAGTGAAGCCCGGACGCTCAGGACCAAGCC 5181
 QY 4918 ATGATCACTCTCTGACTGGAGAGTGAAGTCAAGATGTGGCACTAGTGAAGAACCAAG 4977
 Db 5182 ATGATCACTCTCTGACTGGAGAGTGAAGTCAAGATGTGGCACTAGTGAAGAACCAAG 5241
 QY 4978 CACGAGACCAAGAGAGAGGAGGAGCCGGGAGCAAGATGTGTCTGAATTTACTGACC 5037
 Db 5242 CACCTGAGCAACAGTGAAGGAGCCGCTGAGCAAGAGTGTCTGAGATCTACTTAAAC 5301
 QY 5038 CGACTCTGAGCACTAAGGAGCACTGAGAGAAATTTGATGAGCTCTTGAAGCAATC 5097
 Db 5302 AGGCTGTGAGCCACCAAGGAGCACTAAGAAATTTGATGAGCTCTTGAAGCAATC 5361
 QY 5098 TTGAGACGAGCAACCGTGTCTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 5157
 Db 5362 TTGAGACGAGCAACCGTGTCTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 5421
 QY 5158 CTGAGTGAAGCAGGCTGATTAACATGCAATTCATGACCGGCAAGTCCGCTTACCTGAG 5217
 Db 5422 TTGATGAGCAGGCGCAGCAGATCAGATCCAGATCCAGATGTCGCTTACCTGAG 5481
 QY 5218 AGCAATTTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 5277
 Db 5482 AGCAATTTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 5541
 QY 5278 GACATCCAAAGAGAGATCAAGAGCGCTGCTCTGAGTGGTGGTCAAGCTTCATG 5337
 Db 5542 GACATCCAAAGAGAGATCAAGAGCGCTGCTCTGAGTGGTGGTCAAGCTTCATG 5601
 QY 5338 GACTCTTGTCTCAAGCTGAGAGCAGGCTGGGCAAGAGATGCTGCTCAAGAGCTGTG 5397
 Db 5602 GACTCTTGTCTCAAGCTGAGAGCAGGCTGGGCAAGAGATGCTGCTCAAGAGCTGTG 5661
 QY 5398 TTGCGCAAGAGATCCCGCAAGAGATTGGGTGAGAGGTATTAATCAAGCAATGAGG 5457
 Db 5662 TTGCGCAAGAGATCCCGCAAGAGATTGGGTGAGAGGTATTAATCAAGCAATGAGG 5721
 QY 5458 AAGATTCAGCATCAAGCAGCAAGACATGAACGATACCTGTGAGCAGTCCGAGATG 5517
 Db 5722 AAGATTCAGCATCAAGCAGCAAGACATGAACGATACCTGTGAGCAGTCCGAGATG 5781
 QY 5518 CACATGAATGAGTTCACACATGAGTGAACCTGAGAGATCTTCTCTATGTGGGCAA 5577

Db 5782 CATCTCAGCCAGTTTAAACAGATGAGTGCCTTGCACGAGATCTACTCTTACATGCCCAAG 5841
 QY 5578 TACAGGAGAGATCTCTTGAACCTTGTGACCAAGATGACAGTGTGGAGCAAGATG 5637
 Db 5842 TACAAGAGAGATCTCTGAGTGTGCTTGAAGAAAGATGAAACAGGCTCGAAGAGAGGCTT 5901
 QY 5638 GCTTCAACTGAGACAGATGATTAACCTTCACTGAGCTTGAAGCTGAG 5685
 Db 5902 CGAAGCAACTGAGAGAGTGTGAGACACATGAGCTTGAAGAGCTGGA 5949
 RESULT 10
 HSBXGENE
 LOCUS
 DEFINITION
 ACCESSION
 VERSION
 KEYWORDS
 SOURCE
 ORGANISM
 REFERENCE
 AUTHORS
 TITLE
 JOURNAL
 REFERENCE
 AUTHORS
 TITLE
 JOURNAL
 COMMENT
 FEATURES
 source
 1. 6039
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /chromosome="X"
 /map="q28"
 /feature_type="fetal brain, skeletal muscle"
 /clone_id="lambda ZAP, lambda GT10"
 /gene="sex"
 /gene="sex"
 /gene="sex"
 /note="kozak consensus"
 185..5800
 /gene="sex"
 /codon_start=1
 /product="SEX protein"
 /protein_id="CA61132.1"
 /db_xref="GI:1200235"
 /db_xref="GOA:P51805"
 /db_xref="SWISS-PROT:P51805"
 /translation="MPSVCLLLLLFLAVGALGNRPRAFFVYVDTTLTHAVRYTGE
 VFGAVNRVFKLAPNITERRAVTGVEDNAGALNRPRAFFVYVDTTLTHAVRYTGE
 DVAARLVACGSIWGTGQFLKDDFKGGEFHKKHKLISAGSDPSAKAYIVDQ
 GBEKLFVGAIVDGSFVPTLSRKISIDSDADMSLVYQDFVSQIKTIPSIDLSL
 YPALDIYXYIVGVASFYFTLQDIDTQTLDTGGEFTSKIVKMGDSDFSVY
 EPIGCSWGEVRLVQSAHLAKPGLLAQALGVADBDVLTFTPSQGNKASPPRO
 TILCFTLTNINAHIRRIOSCYRGSTIALPMLINKELICINTPQIKNFSGIYV
 OPIQGLHVEGLPLADSTDMASVAAYVROHSVYFICTRGSLKIVYDGFQDHL
 YETVPPVDSGPILRLDLPDRHHTILSEKVSQIIPVETGROYSCAALGSDPHC
 GNCVLAHRCRGACICGASAPRPAHELSKIQVRYRPNNVSTSGVQLTYTHIVP
 DISAGVSCFEAAENEAVALPSGELCSPSLOERHALTRHGATRYRLQILSET
 GVFAGADVFFNCSVLOSQMSQVSPYCHCKRHTLTSRPHSCFQEGHSHFEG
 CEILPSGDLIPVGMQPLTIRAKLPQPSQGNKYECEVAVQGNQKQVPAVRNFS
 SVQCNASYSYEDGHELDLFSVMDPFIIDKPRFALLVYKMAORPSCGLCLK
 ADRFVNGMCISEHRCQLRTHCPAPRTNMHLSQKTRCSHRIQIHPLVGPKEGT

RTTIVGENILLSREVGIRVAVRNCNIPAEITSAERIVCEMESLVPSPPPGVELC
 VDDCADFERQSEQVSEFTPTFDQVSPBRGASGTLTISGSLDAGSVTVLVRD
 SSCQFBRDAPAKIVICISPLSTSPQAPITLIDRANLSPGLIYTVYODPTVRLB
 TMSIINGSTAITVSGTHLTVQEPVRAKRIETNTQVINDNAMCKAKGFLGR
 POPRAQHEHDEFFGLDHDVQEPVRAKRIETNTQVINDNAMCKAKGFLGR
 KCKNLI PAAGSSRLNLYTLIGGQCSLTPSDTOLCGSPQGTQOPVMTXVGLBEM
 IGLTHISABRLLTPAMWGLAGGLILLAIYAVIYAKRTOPDRITKXJLOJOMN
 LESRYALECKEAPAELOTDINELTNHMEVOJPELDYTYAVRVLPETIEHVPKEL
 DTPVNERKALRLEFGULHSRAVLTHLEAQSSFSMRDQSTVASLTMVALOSRLDY
 ATGLLKQOLADLIERNLESKNHPLKLRTESEVAKMLTMMFTLHFKFLKAGEPL
 FLIYCAIKQOMERKPIDATYGEARYSLEDKLIROQIDYKTLTLHVCPENEGSAQV
 VVINDCSITQAKDLIDTVYKGI PYSORPKAEDMDLEMPQCMTRILIOBEDVYTK
 ECDKRLNLSIAHYOVDSLVALVPEKQSAVYMANSPFTSGLSYESLSTKAGSPS
 LSRBAPMTIPDOETGTMHLYVKNNDHDPBPDGSGPMVESEIYTLTLAKRGLQK
 VDDLEPVTSTHRSALPLAKMFPDLDEADROQISDVPVHTKSKNCLPLRFVY
 NYTKNPFQVFDHKNISIDQACISVAQFPMDSCTSERLISGLSPSNLTLAKQIPVY
 KSWERYRIDAKMASISDQMDAYLVQSRHLHADFVSLALNELYFYVTKYQOELL
 TALDRBASCRKHKLRQKLEQIISLVSSDS"

ORIGIN

Query Match 44.6%; Score 2535.4; DB 9; Length 6039;
 Best Local Similarity 67.3%; Pred. No. 0;
 Matches 3730; Conservative 0; Mismatches 1777; Indels 39; Gaps 9;

QY	152	TCAATCACCTGTGTGATGAGAGACAGACATTTACTTGGGGGCGCTCAATCGGA	211
DB	282	TTACCACTGCTGTGACCGGGTGAAGTGTGTGGGCGCAATGAAACCGAG	341
QY	212	TTTACAGCTCCGACCGACCTGGAAGTCTTGTAACGATGACAGGCGGACGAG	271
DB	342	TTTTAAGGTGGCCCCCACTGACTGAGCTGGGGCCCATATGCAAGAGCCCGTGACG	401
QY	272	ACAACCCCAAGTGTATACCAACCCCGCATGTCACACCTGCAATGAGCCCTGACCA	331
DB	402	ACAAGCTGCTGCTACACCGCCCGCATGTCACACCTGCAATGAGCCCTGACCA	461
QY	332	CCAACATGTCAACAGATGCTCTCATAGACTACAGAGAAACAGGCTGATGCTGTG	391
DB	462	TGGACAACTCAACAGAGCTGCTCATAGACTACAGAGAAACAGGCTGATGCTGTG	521
QY	392	GGAGCCTGTACCAAGGATCTCAAGCTGCTGAGGCTGAGAGACCTCTTCAAGCTGGGG	451
DB	522	GGAGCCTGTGAGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG	581
QY	452	AGCTTTATCATAGAGAGAGACTATCTGTGACGCTGTCACAGAGCGGCTCACTTTG	511
DB	582	AGCTTTATCATAGAGAGAGACTATCTGTGACGCTGTCACAGAGCGGCTCACTTTG	641
QY	512	GAGTATGCTGCTCAAGAGACCTGAGATGACAACTGCTCATTTGGCCAGGCAATGGA	571
DB	642	GAGTATGCTGAGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG	701
QY	572	GGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG	631
DB	702	GGAAGTGGAGAGTCTCCCACTTGAAGCTCCCGCAAGCTCATAGTATGAAAGACAG	761
QY	632	ATGGGATGTCGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG	691
DB	762	GGAAGTGGAGAGTCTCCCACTTGAAGCTCCCGCAAGCTCATAGTATGAAAGACAG	821
QY	692	CGAGACCTTCACTCACTCTGACTTGAATATCATATGATCTATGATTTAGAGAGT	751
DB	822	CAGACAGGCTGCTTGAACCTGCTTGAATATCATATGATCTATGATTTAGAGAGT	881
QY	752	GCAATTTGTCTATTTTGAACCTTGAACCTGAGATGCTTTCACAGGCTTCACCA	811
DB	882	CTCTCTGTGTACTTCTGAGCTGAGCTGAGACCCAGACAGCTGTGTTGAGACAG	941
QY	812	CCAGAGAGAGGATGATATCATATGATCTGAGGCTTTCAGAGAGAGACAGCTTCA	871
DB	942	CGGAGAGAGAAATTTTCAAGTCTCAAGATGCTGAGGAGGAGAGACTCAAGATCT	1001
QY	872	ACTCTATGTAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG	931

DB	1002	ACTCATACGGAATTCCTCCATCGGCTGCTCTGAGGCGGCGCTGAGATACCGCTTGCTG	1061
QY	932	AGGCTGCTTACCTGTCACCAAGCGGAGGCGGCTGCTGAGAGAGACCTTGAAGTCCATCAG	991
DB	1062	AGAGCGGCCACCTGCGCAACCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG	1121
QY	992	ATGATGACCTGCTCTTCAACCGCTTCTTCCAGAGGCGCAAGACGGAATATCCCTG	1051
DB	1122	ATGAGAGCTGCTCTTCAACCGCTTCTTCCAGAGGCGCAAGACGGAATATCCCTG	1181
QY	1052	ATGAGTGGGCGCTGCTGCTTCAATGAGAGATTAATGACCGCATTTAAGAGCGAG	1111
DB	1182	GGGAGACCATCTCTGCTCTTCAACCGCTTCTTCCAGAGGCGCAAGACGGAATATCCCTG	1241
QY	1112	TGAGTCTGTTACCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG	1171
DB	1242	TCCAGTCTGCTCTATGCTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG	1301
QY	1172	TCCGCTGAGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG	1231
DB	1302	TGCTGCTGATCAACACCCCATACAGATCAACCGCACTTCTGTGGCTGTGTTAACC	1361
QY	1232	CTCCCTGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG	1291
DB	1362	AGCTTGGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG	1421
QY	1292	GCAATGCTGCTCATATGCAATATGCTTACAGAACCACTCTCTGAGCTTTGTGGGACCA	1351
DB	1422	GCAATGCTGCTCATATGCAATATGCTTACAGAACCACTCTCTGAGCTTTGTGGGACCA	1481
QY	1352	AAAGTGCAAGCTGAGAGAGAGATCCGGGTGATGAGACCCAGGGGCAACGCTTCAATATG	1411
DB	1482	GCAAGGCAAGCTTGAAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG	1538
QY	1412	AGAGGCTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG	1471
DB	1538	AGAGGCTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG	1598
QY	1472	ACGAGCACTCTCATATGCTGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG	1531
DB	1598	ACGAGCACTCTCATATGCTGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG	1658
QY	1532	GTCATATCAGAGCTGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG	1591
DB	1658	AGCAGTACCAAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG	1718
QY	1592	TGCTGCAACAACGCTGCAACCGGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG	1651
DB	1718	TGCTGCAACAACGCTGCAACCGGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG	1778
QY	1652	TTGCTTGGAGATGAGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG	1711
DB	1778	TTGCTTGGAGATGAGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG	1838
QY	1712	CTCAGTACCAACGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG	1771
DB	1838	CTCAGTACCAACGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG	1898
QY	1772	TCAACTGCACTTTGAGACCTGTGAGAGATGAGATGGG--CTGTGCTGAGGCAATCAGA	1828
DB	1898	TGAGTGGCGCTTTCAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG	1958
QY	1828	TCCAGTGTACTTCCCTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG	1888
DB	1958	TGCTTGGCGCTTTCAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG	2018
QY	1888	ACCATGCTGTAAGCTTCACTCAATCAAGAGAGCCGCGATGACCTTTCGACAGACCA	1948
DB	2018	CCGCACTGTGGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG	2078
QY	1948	GCTTGTCTTTCAATTTGAGGCTGCAATTTGCTGCTGCTGCTGCTGCTGCTGCTGCT	2008

Db 2079 ACTTGTCTCTTCAACAATGACAGCGTCTCCAGTGTGCAATGTCCTGTGTTGGACGCCCTT 2138
QY ACCGTGTGCACTGTGTAAATACCAGGCAATGTCTGACCAATGACCCCAAGACCTGTCTCT 2068
Db 2139 ACCCTGTGCACTGTGTAAATACCAGGCAATGTCTGACCAATGACCCCAAGACCTGTCTCT 2198
QY TCCAGAAAGGCCGAGTGAAGTGTGCCCGAGACTGTCCCGACGCTCTGTGAGTGAACAAGA 2128
Db 2199 TCCAGAAAGGCCGAGTGTGCCCGAGACTGTCCCGACGCTCTGTGAGTGAACAAGA 2258
QY TCCGTGTGCGCGTGGAGGTGATCAAGCTTATCAGCTGAAAGGCCAAGACCTCCCCAGC 2188
Db 2259 TCCGTATCCCGTGTGGGTTCATGACGCTTCTTACCTTGGGGCTTAAGAACTTACTAGC 2218
QY CCCAGTCTGGGAGCGGTGCTTACGAATGATCTCCAACTTCAGGTCAGGCGCAGAGCTGAG 2248
Db 2319 CGCAATCGGGCCAGAAAGAACTATGATGTGCTGTGCGGGTGCAGAGGGCGGAGAGGGG 2378
QY TGCCCGCGCTGTGCTTCAACAGCTCCAGCTCAAGTGCAGAGCAACCTCTTATCTATG 2308
Db 2379 TGCCGTGCGGTGCTTCAACAGCAAGATGTGCAATGTCAGAAAGCCTCTCTTATG 2438
QY AAGGATGAGATGATGATGACACCGAGCTGGAATCTCCGTGTGTCTGGGATGGAATCTCC 2498
Db 2439 ACATTGACAAACCCAGCTCAAGATTAAGTTCACTCTTCAAGTGTGAGCCATGCGTGA 2428
QY GCTGCGGCTGTGCTCAAGCTGACCCAGACTTGCATGTGTGCTGTGCGAGGCCAG 2488
Db 2499 CCATGACAAAGCTCCAGCTTCGAGCTCGAGCTGTGTACAAAGTGTGGGCGAGAGGCCCA 2558
QY GCTGTGCGCTGTGCTCAAGCTGATGCCCGCTTCAATGTGTGTGTGATCTCAAGAC 2618
Db 2559 GCTGTGCGCTGTGCTCAAGCTGATGCCCGCTTCAATGTGTGTGTGATCTCAAGAC 2618
QY GCGATGTCACTCTGTGCGCACTGTCTGCTGCGCAAGAGCAAGTGTGTGATGTGTG 2548
Db 2619 ACAGGTGCGAGCTGTGGAACCACTGCTGCGGCGCCGAGAACCACTGTGTGACTGTGACC 2678
QY GTGCCAAAGCAAGTGTGACAAACCCCGCATCAAGATTAATCCCGGTGACAGGCCCC 2608
Db 2679 AGAAGGCAACCGGTGTGACCAACCCCGCATCAAGATTAATCCCGGTGACAGGCCCC 2738
QY GGGAGGGGGGCAACAAGTCAATTCGAGGGGAGAACCTGTGGCTGTGAATTTGGCA 2668
Db 2739 AGAAGGAGGCAACCGGTGTGACCAATTCGAGGGGAGAACCTGTGGCTGTGAATTTGGCA 2798
QY TGCGCTCCCATGTCAAGTGTGTGAGTGTGAGTGTGAGCTTGTATGTGATGTATCC 2728
Db 2799 TGCGC-----CTGCGGGTGTGCTGTGCTGTGCACTTCATTCGCGCCAGTACATCA 2852
QY CTGCAAGACAGATCGTGTGTGATGAGTGGGGAGGC---CAAGCCCAAGCATGAGCT 2785
Db 2859 GTGTGAGAGGATCGTGTGTGATGAGTGGAGAGTGTGCTGTGCTGCGCCAGCGCGCGGGGC 2912
QY TCGTGAAGTTCGCGTGTGTGTGTGCGCTGTGAATTCATGCGCGCGCTCACAGCTCT 2845
Db 2913 CCGTGAAGT 2972
QY ATTAATTCAATGACATGATCTCTCTCAATCTGAATCTGAACCCGAGGCGGCGCATGTCCGAG 2905
Db 2973 AAGCTTTGTGACCCCAAGTGTGACCAAGTGTGACCAAGTGTGACCCGCGCGCTCCGGG 3032
QY GGAACCAAGTGAATCAATCAAGGCAACCACTGAAAGCCGGAAGCAAGTGTGTGTGTGT 2965
Db 3033 GGAACGAGTTCATCTAGGAGCTCTGTGATGTGGGAGGAGGAGGAGTGTGTGTGTGTGT 3092
QY TTGGAAGAGAGCGCTGTCTCTTCAAGAGCATCTCAATCTCAATTTGTGTGACACCA 3035
Db 3093 TTAGGAGAGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 3152
QY CATCTCAGATGAGTGC---TAGAGATGAAGGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 3082
Db 3153 CTCTCTCAACCTGTGGGCGCCAGGCGCCCATCACTTGTGCAATGACCGGGCTTAACA 3212

QY 3083 T---CCACGAGACCTGTGCTTTCATGATGTGAAGACCCCAACATGCTGTGAGATGAGC 3139
Db 3213 TCTCAGCGCCCGGCTGTATCTTACACTCACTGAGAGACCCACGCTACCGCTTGAGC 3272
QY 3140 CAGAAATGAGCATGTGTGAGTGAACAACACCCATGCGCGTGTGGGGAGCCACCTGAGC 3199
Db 3273 CCACTGTGAGCATCTCAATGTGAAGCACTGCACTATCTGTAGTGGAGCCACCTGTGCA 3332
QY 3200 TCATAAGAAACCCGAGATCGTGTGCAAGCATGAGGAGAGAGCAATCAATATCTGTG 3259
Db 3333 CGATCAGAGAGCCCGGCTCGTGTGCAAGTACCGCGGATTTGAGACCAATATATCTCC 3392
QY 3260 AGTTCTGAAGCTTCTAGATGACCTGTGAGGAGCGCGCGCTGTGTGTGTGTGTGTGTGT 3319
Db 3393 AAGTATCAAGCACTGTGCACTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 3452
QY 3320 ACCAGTCAACCTGACCGAGAGGCGGAGAGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 3379
Db 3453 AGCTGTGAGGCGAAGCGAGCACCTGATGATGTGTGTGTGTGTGTGTGTGTGTGTGTGT 3512
QY 3380 CCTGTCTATCTTCAAGAAAGCAACTTCACTATATCCCAACCGGATTTGAGGCT 3439
Db 3513 CGGCGGCTCTTCAACCGCTCTCTTCTTACTTACTTACTTACTTACTTACTTACTTACTT 3572
QY 3440 TTGATCTCTCAGAAATCTGTGAGCTCAAGCTGTGACAGCCCATCATCTTAAAGGCAAGA 3499
Db 3573 TGGGGCGCTCTGT 3632
QY 3500 ACTGTATCCCGCTGT 3559
Db 3633 ACTGTAT---TCCGCGGAGCGGAGAGCTCCGCTCACTCACTGTGTGTGTGTGTGTGT 3685
QY 3560 AGAAGCGGTGACCGGTGACCGTGTGATGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 3619
Db 3690 GCGAGCGGT 3749
QY 3620 TCGGAGGCAAAAGT 3679
Db 3750 CTGGCGGAGCGCTGT 3809
QY 3680 ACATTCGCCCGGAGACCGCTCAAGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 3739
Db 3810 ACATTCGGGAGAGCGGAGCGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 3869
QY 3740 GCTTCTCATCTTTTCAATGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 3799
Db 3870 GCTTCTGT 3929
QY 3800 GTTACTCAAGCTGAAAGCGGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 3859
Db 3930 CGGACCGT 3989
QY 3860 TGTGTGTGAGAGAGCTTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 3919
Db 3990 TGTGTGTGAGAGAGCTTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 4049
QY 3920 TGTGTGTGAGAGCGGAGCTTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 3979
Db 4050 TGTGTGTGAGAGAGCTTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 4109
QY 3980 CAGAAATTTGAAGACCACTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 4039
Db 4110 CGGAGATGAGAGCGGAGCTTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 4159
QY 4040 GTGTGTGAGAAAGCGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 4099
Db 4160 --GTGAGAAAGCGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 4217
QY 4100 CTTTCAATCGCAGCTTGT 4159
Db 4218 CTTTCAATCGCAGCTTGT 4277

OY	4160	CTCACTCATATATACCGTGTGTGAGAGCAAGCTGAGATGAGCCACTGATGTGTGAGC	42123
Dp	4278	CTTGCTGTACCATATGTGTGGCCCTGTGACAGCCGGCTGTACTATGTGCCACGGGGCTGTCAAGC	4337
OY	4220	AGCTGTGGCCGCACTCTATTGACAAAGAACTGTGAGAGCAAGAACCACTTAAGCTGTGC	4279
Dp	4338	AATGTGTGGCCGCACTGTATGTGAAGAAACCTGTGAAGAAACACACCCCAAGTGTGTGC	4397
OY	4280	TCAGAGGACTGATCAATGTGGCTGTGAAGAAAGTGTGACCAATTGGTTTACTTCTCTCT	4339
Dp	4398	TACCAAGGACAGAGTCATGTGGCTGTGAAGAAAGCTTACCAACTGTGTCAAGTTCCTGTGC	4457
OY	4340	ACAAGTTCCTCAAGAGAGTGTGTGGGGAGCCCTCTTCTCCCTGTTCGTGCCATCAAGC	4399
Dp	4458	ATTAAGTTCCTAAGAGAGTGTGTGGGGAGCCCTCTCTCTCTTACTGTGTGCATCAAGC	4517
OY	4400	AGCAGATGAGAAAGGGCCCAATTGACCGCATACGGGGAGGGCCCGCTACTCTTGAAGC	4459
Dp	4518	AGCAGATGAGAAAGGGCCCAATTATGTGCATACGGGGAGGGCAACGATACTCCTTGAGC	4577
OY	4460	AGGACAACTCATCGCGCAGAGATGTGATCAAAACCTGTGTCTGTAGCTGTGTCAAGC	4519
Dp	4578	AGGACAACTCATCGGTGTAGAGATCGACTAAGAACACTGACCTTTCATCTGTGTGTCTC	4637
OY	4520	CAGACAAATGCCAACAGCCCGCAGGTCCCGATAAAGATCTTCACTGTGTACACATCACTC	4579
Dp	4638	CGAGAAAGAGGGGAGCGGCCCAAGTCCAGTTCAGTGAAGGTTCTCACTGTGTACAGATCAACC	4697
OY	4560	AGGTCTAAGAGAAAGATTTGTGATGCACTTTCAAGAAATGTGCCTTGTCTCCACCGGCCCA	4639
Dp	4698	AGGCAAAAGATTAAGCTGTGTGACACTGTGTACAAAGGGCATTCGATCTCCAGGTGCCA	4757
OY	4640	AAAGCTGCAGATATGATATGTGAGATGTGCGCACAAAGAAAGTGGGCAAGATGATCTTGTGAGG	4699
Dp	4758	AAAGCTGAGGATATGATCCTGTGAGATGTGCGGCCAGGGCCGATATCTCGATATCTTCCAGG	4817
OY	4700	ATGAAGACATACCAACCAAGTTTGAATGATTTGAAAGCGACTGACACACTGTGCCACT	4759
Dp	4818	ATGAGGATGTACCAACCAAGATCGAGTGTGACTGGAAAGGCTCAACTCACTGTGCCACT	4877
OY	4760	ACCAAGTCTCAGATGATGTTCGGTGGGGGATTAAGTCCAAACAGGTGACAGCTATTAAGC	4819
Dp	4878	ACCAAGTGTACAGACGGTTCCTTGTGTGAGATGTGTGCGCCAAACAGTGTGTGCTATTAACA	4937
OY	4820	CAGTGAACAACATCCACCGTGTCTCAAGGACTGTGACCAAGTAAATATGAACAATGATCCGCT	4879
Dp	4938	TGGCCAACTCTTCACTTCAACCGS---CTCCCTCAAGCGGCTGTGAGAGCTGTCTCGCA	4994
OY	4880	ACAAGGGAGAGCCCGACAGCCTCGGCTACAGGACACCTATATATCACTCTGTACCTTGAGA	4939
Dp	4995	CGGCGAGAGCCTGTATAGCTTCGGCTACGGGCAACCATATATTAAGCTGTACCAAGAGA	5054
OY	4940	GTGAGATCAAAATGTGTGCACTTAGTGAAGAACCAAGACAGGAGACCAAGAGAGGGGG	4999
Dp	5055	CAGGACCAAAATGTGTGGCACTGTGTGAAGAAACCAAGACCAATGCGGACCATGTGAGGGGG	5114
OY	5000	ACCGGGGAGCAAAATGTGTGTCTGAATCTTACCTGTGACCCGACTCTGTGCACCTTAAGGGCA	5059
Dp	5115	ACCGTGGAGAGAAATGTGTGTCTCGAGACTTACCTGTGACACGGCTCTGTGCCACCAAGGGCA	5174
OY	5060	CAGTGCAGAAATTTGTGTATCACTCTTTGAGACCATTTTCAAGACAGGCAACAGCTGTGCT	5119
Dp	5175	CAGTGCAGAAATTTGTGTATGATCCTTTTGAAGACAGTGTTAAGACACGCCACCGGGGCT	5234
OY	5120	CTGCCCTTGCCCTGTGCCATCAAGTACATGTTTGACTTCTGTGATGAGCAGGCTATTAAC	5179
Dp	5235	CGGCTCTTGCCCTGTGCCATCAAGTATCAATGTTTGAATCTTCTGTGATGAGCAGGCAACAGC	5294
OY	5180	ATGGCATTCATGACCCCGACGCTCGGCACTACTGTGGAAGCAATGTGCTGCGCTGTAGGT	5239
Dp	5295	GCCAGATCAGGAGACCCCGATGTGTGCGCACACTGTGGAAGCAACTGTGCTGCGCTGTGACT	5354
OY	5240	TTTGGGTCAACATGATCAAGAACCGGCAATTTGTGTGTGATCATTCATTAAGAACGATCA	5299

Db	Sequence	Score	DB	Length	Matches	Conservative	Mismatches	Indels	Gaps
Db	5355 TCTGGGTAATGTATCAAGAACCCCGAGTTCTGTTTCACATCCAAAGAACGACATCA	44.5%	5355	5895	0	1712	81	17	0
Qy	5300 CAGACGCGCTGCTCTCTGTGTGTGTGCAGACCTTCATGACCTCTTGTCTCCAGCTGAGAGC	44.5%	5300	5895	0	1712	81	17	0
Db	5415 CGATGCTGCTGCTCGGTGGTGTAGCCAGACCTTCATGACCTCTGCTCTACATCCGAGC	44.5%	5415	5895	0	1712	81	17	0
Qy	5360 ACCGGCTGGGCAAGAGATCGCCCTCCCAACAGCTGCTGTATGTCACAGGATATCCCACT	44.5%	5360	5895	0	1712	81	17	0
Db	5475 ACCGCTGGGGAAAGACTCGCCCTCCCAACAACTGCTCTAGCGCAAGAGACATCCCACT	44.5%	5475	5895	0	1712	81	17	0
Qy	5420 ACAAGATTGGGTGAGAGGATTTACTCAGACATAGGAAAGATGCCAGCCATGACGAGC	44.5%	5420	5895	0	1712	81	17	0
Db	5535 ACAAGACTGGGTGAGAGGATTTATGAGACATGCAAGATGCAATGCATCAGGAGC	44.5%	5535	5895	0	1712	81	17	0
Qy	5480 AAGCATGAACGCACTACCTGGCTGAGACATGCCGATGCAATGATGAGTTCAACACCA	44.5%	5480	5895	0	1712	81	17	0
Db	5595 AGACATAGAGATCCCTACCTGGTGGAGAGTCGCCCTCCACGCCAGGACCTTAGCTCC	44.5%	5595	5895	0	1712	81	17	0
Qy	5540 TGAATGCACTCTCAGAGATCTTCTCCATGTGGGCAATTCAGAGAGAGATCCTTGGAC	44.5%	5540	5895	0	1712	81	17	0
Db	5655 TGAATGGGCTCAAGAGACTGATTTCTATGTCAACAAAGTACCGCAGAGATTTCTACGG	44.5%	5655	5895	0	1712	81	17	0
Qy	5600 CTCTGACCAAGATGACCAATGTGTGGAGCAAACTGSCCTTCAAACTTGAAACAATCA	44.5%	5600	5895	0	1712	81	17	0
Db	5715 CTCTGACCAAGATGACCAATGTGTGGAGCAAACTGSCCTTCAAACTTGAAACAATCA	44.5%	5715	5895	0	1712	81	17	0
Qy	5660 TAACCTCATGAGCTTAGACAGCTGA 5665	44.5%	5660	5895	0	1712	81	17	0
Db	5775 TCAGCTGTGTCTCAGCAAGACTTAA 5800	44.5%	5775	5895	0	1712	81	17	0

Db 337 CCTGTGAGGACAAAGAGTGTACCCGCGCCGAGCGTGTAGTCTGCCCCACGG 336
QY 321 CCTAGCACCAACCAATGTCAACAAGATGCTCTATAGACTACAGAGAAAGAGCT 380
Db 397 CCTGGGAGTACTGACAAAGTCAACAGAGTGTCTGTCTGTGAGTATCCGCTAACCGCT 456
QY 381 GATTGCTGTGGAGCTGTACCAAGGATTTGCAAGCTGTGAGGCTGAGAGCTCTT 440
Db 457 GCTGCTGTGTGACGCGCTCCCAAGGAGTGTGCAAGTCTGTGCTGTGACAGATCTCTT 516
QY 441 CAAGTGGGGAGGCTTATCAATAAGAGAGCACTATCTGTCAAGTGTCAACAGAGAGCG 500
Db 517 CAACTGGGTGAGCCACACACCGTMAAGAGCACTACTGTCCAGCGTTCAGAGAGAGG 576
QY 501 CTCACTCTTTGAGTGTGTCTCTTACA-----GCAACTGTGATGACAACTCTTCA 554
Db 577 CAGCATGGCGGGCGTGTCTATTGCGGGCCACCGGGCCAGGGCCAGGCTCTTCTGT 636
QY 555 TGCCACGGCAGTGGATGGGAAAGCCGAGTATTTCCACCATCTCCAGCGGAAACTGAC 614
Db 637 GGGACACCCATCGATGGCAAGTCCGAGTACTTCCCACTGTCCAGCGCTCGGCTCAT 636
QY 615 CAGAAGCTCTGAGCGGATGGCATGTTGCGGTACGCTCTTCCATGATGAGTTCTGTGCGCTC 674
Db 697 GGGCAACGAGGAGGATGGCCGACATGTTGCGCTGTGTGATCCAGGATGAGTTGTGTATC 756
QY 675 GATGATTAAGATCCCTTGGGACCTTCACCATCATCGCTGACCTTGTATATCTACTATGT 724
Db 757 ACAGTCAAGATCCCTTGGACGCTGTCCAGTTCGCGCTTGTGACATCTTATATGT 816
QY 735 CTATGTTTGTAGAGTGGCACTTGTCTACTTTTTCACCTCCACCTGATGATGTGT 794
Db 817 GTACAGCTTCGCGAGCGAGAGTTGTCTACTACCTCAGCTGAGCTAGACACAGCT 876
QY 795 TCCACAGGCTTCACACACAGAGAGAGGTATATACATCCAGCTGTGAGGCTTTGCCA 854
Db 877 GACCTCGCTGATGCGCGCGGAGCACTTCTTACGTCMAAGATCGTGGCGCTGTGT 936
QY 855 GGAGAGACAGGCTTCACTCTTATGTAGAGTGGCCATTGGCTGTGAGCGCAGTGGAGT 914
Db 937 GGAAGACCCCAATTCTACTGTGATGATGCCCATTTGGCTGCGAGCGAGGCTGT 996
QY 915 GAGTATCGGCTGTGCGAGGCTGTACTTGTCCAAAGCGGGGCGGTGTCTGCAAGAC 974
Db 997 GAGTATCGGCTGTGCGAGGATGTCTTACTGAGCGGCGCGGCTGTGCGCCACCA 1056
QY 975 CTTGAGTGCATCCAGATGATGACTGTCTCTTCAACGCTTCTTCAAGGCGCAGAGCG 1034
Db 1057 GTTGGGCTGTGTGAGAGAGAGAGTGTGTCTGTCTGTCTGTCTGTCTGTCTGTCTGT 1116
QY 1035 GAAAAATGAATCTGTGATGAGTGGCGCTGTGTCTTCTTGTGAAGCAGATTAATGA 1094
Db 1117 CCGCGTGAAGCCACCAAGAGTCAAGCTGTGCTGTGTCAAGGCTCAAGGCGCATCAAGA 1176
QY 1095 CCGCATTAAGAGCGGCTGTGAGTCTTGTACCGGGGCGAGGCGACGCTGAGCGCTG 1154
Db 1177 GAAAGTTAAGAGCGGCTGTGAGTCTTGTCTGTGTGTGTGTGTGTGTGTGTGTGTGT 1236
QY 1155 GCTCAAGGTGAAGCATTCCTGTGAGAGTGGCTTAAACATTTGAGTAATCTTGTG 1214
Db 1237 GCTGTCTCAACAGAGGCTGTGAGTGTGATCACTGCGCTGTGCAAGTGTGAGTGTGTG 1296
QY 1215 TGGCGTGGACATGATGTGCTGTGTGAGTGTGCGAGTGTGAGTGTGAGTGTGAGTGTG 1274
Db 1297 CCGGAGAGCTTCAACAGAGCGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1356
QY 1275 CACGAGAGCAGGAGCGGATGAGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1334
Db 1357 GGTGAGCAAGGATGT 1416
QY 1335 GGCCTTTGTGGGCAACAAAGTGGCAGCTGAAAGAGATCCGGGTGTGTGTGTGTGTGTGT 1390

Db 1417 GGTATTGCGCGACAGGAGTGGCGCGATCCGCAAGTCTGTGTGTGTGTGTGTGTGTGT 1476
QY 1391 -----GGGCAACGCTCTTCAGTATGAGACGGTGTGAGTGTGTGTGTGTGTGTGTGT 1445
Db 1477 CGGTGGCGGCGCTGTGCTGT 1536
QY 1446 CCGGATATGAGCTTCTTCCAGAGAGCAGAGAGCACTATCATCATGTGTGTGTGTGTGTGT 1505
Db 1537 GCGAAGCTCTGTCTTCAAGCGCCCAACCAAGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1596
QY 1506 CACAGAGTCCCTGT 1565
Db 1597 GAGCGGGGTGCTGT 1656
QY 1566 AGGCAACCCCACTGT 1625
Db 1657 ACAGGACCCCACTGT 1716
QY 1626 TGAGCGGTCCAGAGAGCGCCGAGGTTGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1685
Db 1717 TGAGCGAGCAGACGAGCCCAAGCGCTTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1776
QY 1686 GGTCCATCCCAATATCTCTGT 1742
Db 1777 TGTGACGCGCGCAATGT 1836
QY 1743 GTACATGTCCCGAGCTGT 1802
Db 1837 CTGGAAGCGTGTGACCTGT 1896
QY 1803 GGTAGGCTGT 1862
Db 1897 TGAAGCGGT 1956
QY 1863 CCGGATCA-----TCAAGAGATGGGAGCACCAGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1916
Db 1957 GCCATACCGGGGCGCAGGGT 2016
QY 1917 AAGAGACCGGATGACCTTGT 1976
Db 2017 CAGGAGACAGGAGAAAGT 2076
QY 1977 CAATTG---TGCCTGT 2033
Db 2077 CAGTGTGAGT 2136
QY 2034 GATGT 2093
Db 2137 CACGCTGTGACACACAGAGT 2196
QY 2094 CAGAGCTGT 2153
Db 2154 GCTTATCAGCTGT 2213
QY 2197 TGAAGCTGT 2256
Db 2257 ACCCATCACTGT 2316
QY 2214 ATGCACTGT 2273
Db 2317 GT 2376
QY 2274 CAGGATCAGT 2333
Db 2377 CAGCTGTGAGT 2436
QY 2334 CGT 2393
Db 2437 AGTGAACCTGT 2496
QY 2394 AGTTCACTGT 2453
Db 2497 GCGCACCTGT 2556

QY	2454	CCCGAATTGCGCATGTGGCTGATGCGCAAGGCGCCCAAGTGCACCCCTGGCGCCACAGACTG	2513
Db	2557	CCCGCGCTTCGAAAGTCCGGATGTGCGTGGCCGAGCGCCGCTGTCTCCCTCGCAACACACTG	2616
QY	2514	CCCTGCCCCA--GGAGAGCCAGATGGCTGAGCTGTCTGTGTCCCAAAAGCAAGTGCACAA	2570
Db	2517	CGCTCCGACACACACTGTGATGTGGATGACAGCGCGCTCACCGGACAGACTCGCTGCACCGA	2676
QY	2571	CCCCCGCATCACAGAGATATATCCGGGTACAGCGCCCCCGGAGAAAGGGGACCAAGATCAC	2630
Db	2677	CCCCAAGATCTCTCAAGCTGTCTCCCGAGAACGGGCTCCAGAGGACGGGACACGGCGCTCAC	2736
QY	2631	TATCCGAGGGGAGAACCTGGGCTTGGAATTTTCGGACATCGCTTCCTCCATGTCAAGGTTGC	2690
Db	2737	TATCAACGGCGCAGAACCTGGGCTTCGCAATTGAAGACCTGTGCGTGTGGCGCGCGGTGGG	2796
QY	2691	TGGCGTGAATGTCAGCCCTTTAATGTGAATGTATCATCTCTGCAGAACAGATCGTGTGTA	2750
Db	2797	CAAGGTGTGTGCAGCCCTGTGGAGAGGAGATCATCAGTGGGAGAGACATCTGTGTGA	2856
QY	2751	GATGGGGGAGGCGCAAGCCCC--AGCCAGCATGCAAGCTTCGTGAGATCTGCGTGGCTGT	2807
Db	2857	GATGGGGGAGCGCCAGCTCCGTGCGTGGCCCAATGACGCGCTGTGTGAGAGGTGTGTGGGGA	2916
QY	2808	GTCGTGGCTGAATTCAATGAGCCGGTCTCTCAGCTCATTAATTCAATGACACTGACTCT	2867
Db	2917	CTGCTCACACACACTACGGCGCTCTGTCAACCAGGCTTCACCTTTCGTGACACCAACTT	2976
QY	2868	CTCAGATCTGAAGCCCAAGCGGGGGCCCATGTCCGGAGGAGCCCAAGTGAACCATCACAG	2927
Db	2977	CTACCGTGTGAGCCCTCCCTGGGCGCTCTGTCAAGGGGGCACCTGTGAATTGGCATCGAGG	3036
QY	2928	CACCAACCTGATACCGGAGACCAAGTGTGTGATGTGTGAAAGACAGCCCTGTCTT	2987
Db	3037	AAGCACTCTGAACGAGGAGAGTGAATGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT	3096
QY	2988	CCACAGGCGATCTCCATCTCTCATTTGTGTGAAACCAACATCTTCATGATGAGTGTCTAGA	3047
Db	3097	CTCTGTGTCCAGAGAGAACTCCCGTAAGTCCGGTCCGTGACACCCCGGGGACAGAGCC	3156
QY	3048	GATCAAGGTGTCCGTGCGAGGTGGAAGAGGCC-----AAGATCCACAGAGACTGTGT	3098
Db	3157	TGGCAGCGCTCCCATCATCATCAACATACACCGCGCCAGCTCACCAACTGTGAGTGA	3216
QY	3099	CTTTCAATGATGTGAAGAAGCCCAACCATGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT	3158
Db	3217	GTAACAATACACCGAGAGACCCACCATCTGAGATCAACCCCGAGTGACATCAACAG	3276
QY	3159	TGGAAACACACCATTCGCGTATGGGGGAGCCCACTGTGACCTCATACAGAACCCCGCAT	3218
Db	3277	CGGTGGAGACCTCTCTGACGTGCACAGGACCAACCACTGTGCACATGTCTGTGAACCCCGAT	3336
QY	3219	CCGTGCCAAGCATGAGGAGGAAAGAGCAATCAATCTGTGAGGTTCTGAACGCTACTGA	3278
Db	3337	CCGGGCGCAAGTATGAGGAGCAATTGAG--AGGAGAACTGTCTGTGTATCAATATGACACAC	3393
QY	3279	GATGACCTGTAGAGGCGCGCGCTGTGGTCTGTGACCAACAGTACAGACTGTACCGGA	3338
Db	3394	CATGTATATCGCGCGCTCCGTCTGTGGCCAACTGTGTGGCAGCCCAACAGAGCTGGGGGA	3453
QY	3339	GAGGCGCAGAGATTGGCTTTATCTGTGAACAACGTTCAGTCCCTGTCTATCTTCAACA	3399
Db	3454	GCGGCGGATAGCTGTGGCTTGTATATGAGCAACAGTGGCTCTCCGTCTGTGTCAATCTC	3513
QY	3399	GACCAACTTCACCTATATCCCAACCGGATGTTTAAAGCTTTGGTCCCTCAGGAATCTT	3458
Db	3514	CACCTCTCTCTCTACTACCTGTGACCCGTACTGAGAGCACTCAGGCCCACTGTGGCTGT	3573
QY	3459	GAGGCTCAAGCCTGTGCGAGCCCATCATCTTAAAGGGCAAGAACTGTATCCGCTGTGGC	3518
Db	3574	GAGGCTGAAGCCCACTCTCCCACTATCTCTCAAGGGCCGGAACTCTTGGCCACCT--GC	3630

QY	3519	TGGGGGCGAAGTGAAGTGAACCACTGTCGTGTGGGAGAAAGCCGTGCAACCGTAGC	35718
Db	3631	ACCGGCGAATCCCCGACTCAACTCAAGGTGTCTATGGCTCCACACCTCTGACCTCAC	3690
QY	3579	CGTGTCAAGTGTCCAGCTGTCTTCCGAGTCCGCCAACCTCATTCGCGAAGCCAAAGTAT	3638
Db	3691	CGTGTCCGAGACCGCAATGCTGTGTCGAGGCGCCCAACTCATCTGGGCGAGCAAGGTGAC	3750
QY	3639	GGCCCCGTGTGGTGGCATGGAAGTACTCCCGGGGATGATGTATCACTTCCCCGACAGGCC	3698
Db	3751	GGTCCGTCAGAGTGGCTTCAGTTCCTCCGACGGGACACTGAGGTGTACTTGGCAGAGCTT	3810
QY	3699	GCTCAGCGCTACCGGCATTCGTACAGCATCGAGTGGCTGCGGGCCCTCTCATATTTTCAT	3758
Db	3811	GCTGACCTGTCTGCTCATTTGTGGGCATTGGGCGAGGCGGGGGGTCTCTGTGCTGTGCAT	3870
QY	3759	CGTGGCCGTGTCTCATTTGCTATTAACGCAAGTCCCGCGAAATGTACTCAGCTGAAGCG	3818
Db	3871	CGTGGCTGTGTCTCATCTCCCTACAGGCCCAAGTGCAGAGATGTGTACCGGCACCTCAAGCG	3930
QY	3819	GCATGACATGACATGGAGCAACTCGAGAGTCCCGGTGAGCCCTGGAGTGCAGAGAAAGCTT	3878
Db	3931	GCTGCACCTCGAGTGGACAACCTGAGTCCCGGTGGCTTGAAATGCAGAGAAAGCTT	3990
QY	3879	TGCCGACCTGCAAGCGGACATTCATGAGCTGACCTGTGATGAGCTCGGGAATTC	3938
Db	3991	TGCAGACTGCAGACAGACATCCACGAGCTGACCAATGACCTGGACGGTCCGCGATCC	4050
QY	3939	GTTCTCTGCACTATTAAGACTTAAACCAATGCGGGGTGCTGTCCAGAGATTGAAGACACCC	3998
Db	4051	CTTCTCTTACTACCGGACATATGTGCATGACGGGTCTCTTCTGTGGATTCAGGACACCC	4110
QY	3999	TGTCTCTCCGGAACCTTAGGTCCCGGGCTACCGGCAAGAGGTGTGAGAAAGGCTTGA	4058
Db	4111	TGTCTCAAGGAGATGAGGATCAAGGCAA-----TGTGAGAAAGTGTGCTGAC	4158
QY	4059	GCTTCTGCCAGCTCATCAACAAGAGTGTCTGTCTCTTCACTCCGACAGCTTGA	4118
Db	4159	ACTGTTTGCGGAGGTGTGTGACCAAGAGCACTTCTCTGACCTTCACTCCGACAGCTGGA	4218
QY	4119	GTCCACAGTATGCTTCCATGCGGACCGTGGCAAGTGAGCTCCTCATATGACACGT	4178
Db	4219	GGCAGACGCAAGCTTCTCCATGCGGACCGGGAAATGTGGCTCTGCTATCAATGAGCGG	4278
QY	4179	GCTGCAGAGCAAGCTGAGATACCCCACTGATGTGTCTGAAGCACTGCTGCGGCACTCAT	4238
Db	4279	CCTGCAGGCGAGATGGAATACGCCCAAGCGTGTCTCAAGGACGTGTCTTCCGACTCAT	4338
QY	4239	TGAACAACCTGGAGAGCAAGAACCAACCTAAGCTG---CTGCTCAGGAAGCATGAGTC	4298
Db	4339	CGAACAACCTGGAGAGCAAGAACCAACCTAAGCTGTACTGTGCGGCGCAACTGAGTC	4398
QY	4299	AGTGGCTGAGAGAGTGTGCAACAAATGGATTACTTTCTCTCTCAAAAGTTCTCTAAGA	4358
Db	4399	GGTGGCAGAGAGATGTCTACTAATGATTCACCTTCTCTTGTATTAAGTTCTCTCAAGA	4458
QY	4356	GTGTGCTGGGAGGCCCCCTTCTTCTCCCTGTCTGTGCAATCAAGACAGAGATGAGAAAGG	4415
Db	4459	GTTGGCTGGGAGGCGGTGTTCATGTCTGTACTGCGCATCAAGACAGAGATGAGAAAGG	4518
QY	4416	CCCCATTGACGCAATACAGGGCGAGGCGCGTACTCTTGTGACAGAGCAAGCTCATCG	4475
Db	4519	CCCCATTGACGCAATACAGGGTGAAGCAAGCTTACTTCCCTGTATGAGCAAGCTCATCG	4578
QY	4476	CGACGATTTGACTTCAAAACCTGTGTCTGAGCTGTGTCAAGCCCAAGCAATGCCACAG	4533
Db	4579	GCACGATTTGACTTCAAAACCTGTGAACCTGTGAACCTGTGAATGAGAAATGC	4638
QY	4536	CCCCGAGGTCCCAATAAAGATCTCAACTGTGACACCATCACTCAGGTCAAGGAGAAAT	4595
Db	4639	ACCTGAGGTCCGGTGAAGGGGCTGAGCTGTGACACGCTACCCAGGCGCAAGGAAACT	4699
QY	4596	TCTGGATGCCATCTTCAAAATGTGCTTGTCTCCACCGGCGCAAAAGCTGCAGATATGGA	4655

VEDLAGVNCSEFDETESESVLEDRHCHRSAREAVAPITRGQDQRVKLYKKE
 TGRKASVDVFFVNCSSVQSCISCVNGSFCMCKRYACHNVADCFLEGRVNVSE
 DCPQILSPQIVVGVVXPTILAAKNI.PQPOSQGVCEFLH.PGSAPATVAFNS
 SLOOCSNYSYEGNDVSLPVNL.SVNMGNVINDPNIOAHLYKCALRESGGL
 KADPRFCMGCVAERBCSLRHHCADPTASWMAHNSRCPDPLIKLSPEPPOG
 GTRILITGENTLARPEDVRLGTRVQKULCSFRESYITABOIVCEIGASSVRADL
 VAVCVRDSPHYRALSPKSFVITVPTFRVSPRSRPGWGTWILBGLHAGSDVAV
 SVGGPCCSFWSNRSRIRCLTPRQSPSAPILININAQOLNPEKCNVYEDPTILR
 IDBWSINGERGTLVTGTNLAVREPRIRAKYGGIENEGCLVYNDTPWLCRASPVA
 NPYRSPLEGDELFMDVRSLLVNSFSFLYEDPVLPSFTGLLEKPSF
 LILKGNILPAPGMSRLVYTLIGSTPOTLVSFTOLLEAPNLTGQHKTVAGGFL
 FESPTIOYSDSLTLPAIVIGIGGGGGLILVAVLILAKRSRDAFRLKXLO
 MNLEBRVALBECKAPAEIQTIIHNLINDLDAIGIPRIDRYTAMRYV.PRIEDHPIL
 KMEVAVANYSKSLTLEQULTKHFLTIFRTLEAKSF.SMEDRGNVSLIMTALQGS
 MEYATGVLKQLSLDLIERULESKNHPKILRATESVABEKLINFTLVLFLKQGS
 BEPLMYLCAIKQOMKGPIDATGEARVSLSEDKLIRQIDYKILTLNCPVENBP
 EYVKGJLDCDVTQAKERLLDAAYGVFVSPQPKAADMLERQGMARILLOBEDVT
 TKIDNMKRLNLAHOVTDGSSVALVJKOTSANYI.SNSSTKSLSEYSEMLRTAS
 PLSRSRPMITPDL.ESGTKMLHLYKNHNDLHOREGDKWVSLVYLTRILATKGLT
 QKRFVDIPETISTHRSALPLAKYMFDFLBCADHGQIHDAVDRITWMSNCLPLR
 FVNVIKNEQFVFDLHKSLTDACLSV

ORIGIN

Query Match 42.6%; Score 2423.2; DB 9; Length 5263;
 Best Local Similarity 67.7%; Pred. No. 0;
 Matches 3525; Conservative 0; Mismatches 1639; Indels 42; Gaps 8;

QY 141 CGCCGAGGCTTTCATCACCTGCTGATGAGAGACAGACACATTACTGGGGGC 200
 DB 76 CGACTGGGGCTCACACCACCTATTGCTGATGAGAGACAGGCGAGGTATGCTGGGGC 135
 QY 201 CGCATATGGAATTTACAGCTCTCCAGGACCTGGAAGCTTGGAGCGCATGAGAGACAG 260
 DB 136 AGTGAACCCCATTTAAGCTGTGGGGAACCTTGACATGCTGCGGGCCACAGTCAACGG 195
 QY 261 GCCGAGAGAGACCAACCCCAAGTGTACCCACCCCGCATGCTCCAGACCTGCAATGAGCC 320
 DB 196 CCTGTGAGAGACCAAGAGAGGTACTACCGCGCCANCGTGCAGTCTGCCCCCAAG 255
 QY 321 CCTGACCAACCAACAAATGTCAAGATGCTCTCTCATAGACTACAGAGAAACAGGCT 380
 DB 256 CCGGGCAATACAGCAACGTCACAAAGTGTGCTGCTGAGACTATGCCGTTAACCGCT 315
 QY 381 GATTGCTGTGGAGGCTGTACCAAGCATCTGCAAGCTGTGAGAGCTTGAAGACCTCTT 440
 DB 316 GCTGGCTGTGGAGGCTCTCCAGGGCATCTGCGCACTGCTGCGCTGGAACATCTCTT 375
 QY 441 CAAGCTGGGGAGGCTTATCATTAAGAGAGCATATCTGTCAAGGTGTCAAGAGAGGG 500
 DB 376 CAATCTGGGTGAGCCACACACCGTTAAGAGACATACGTCCAGGTGACAGAGAGAGG 435
 QY 501 CTCAAGCTTGTGAGTATGATGCTCTCC-----TACAGCAACTGAGTGAACAAGCTGTGAT 554
 DB 436 CAGCATGGGGGGGTGTCTATTGCGGGCCACCGGGCCAGGGCCAAAGCTCTTGGT 495
 QY 555 TGCACAGGAGTGGAGAGCCCGAGTATTTTCCACCAATCTCCAGCCGGAACCTGAC 614
 DB 496 GGGCAACACCATGATGAGCAAGTCCGAGTACTTCCCAACATGTCAGCGCGTCAAT 555
 QY 615 CAAGAATCTGAGGCGGATGGGATGTTTGGGTAACGCTTCAGATGATGATGCTGGGCTG 674
 DB 556 GGGCAACAGAGAGATGCCACATGTTCCGCTTCGTGACAGAGATGATTTGTGTCAATC 615
 QY 675 GATGATTAAGATCCCTTCGACACCTTCAACCATCATCCCTGACTTGTGATATTAATGT 734
 DB 616 ACAGCTCAAGATCCCTTCGACACGCTGTCAAGTCCCGGCTTGTGACATCTACTATGT 675
 QY 735 CATATGTTTTAGAGTGGGCAACTTGTACTATTTTGAACCTTCAACCTGAGATGCTGTC 794
 DB 676 GTACAGCTTCGAGAGAGAGTGTGTACTACTCTACGCTGACACTGACACACAGT 735
 QY 795 TCCACCAAGGCTCCACCAACCAAGAGAGAGGTGTATACATCCAAAGCTGTGAGGCTTTGCA 854

DB 736 GACCTGCTGTATCCGCGCGGACAGCATTTTCAAGCTCCAAAGATGTGCGGCTGTGTGT 795
 QY 855 GAGAGACACACCTTCAACTCTTATGATGAGAGTGGCCATTTGCTGTGAGCGCAGTGGGT 914
 DB 796 GAGAGACACCAATTTCACTGTATAGTGTAGGTTCCCATTTGCTGTGAGAGACGGGTGT 855
 QY 915 GAGATACCGCTGTGTGAGGCTGCTTACCTGTCCAAAGCGGGGCGGTGTGTGGCAGAGC 974
 DB 856 GGAATACCGGCTGTGTGAGGATGCTTACCTGAGCCGGCGCGGTGCTGCTGCCACCA 915
 QY 975 CTTTGAAGTCCATCCAGATGATGACCTGCTTCAACCGTCTTCTCCAAAGGGCCAGAGCG 1034
 DB 916 GCTGGGCTGTGTGAGAGACAGAGACGTGCTTCAACCGTGTGAGAGGAGAGAA 975
 QY 1035 GAAATTAATTCCTGTGATGAGTGGCGCTGTGCATCTTCACTTTGAGCAGATTAATGA 1094
 DB 976 CCGGCTGAGCCCAACAAAGAGTACAGCATGTGCTCTTCAACGCTCAGGGCCATCAAGGA 1035
 QY 1095 CCGCATTAAGGAGGCGGTGACATCTTGTATCCGGGGGAGAGGACGCTGACCTGGCTG 1154
 DB 1036 GAAATTAAGAGAGCATTCAGTGTCTGTCTACCTGTGTGAGAGGAGAGCTCTCCCTGCGGTG 1095
 QY 1155 GCTTCAGGTGAGAGACATCCCTGACAGAGTGGCTTCTTAACCATTAAGCATTAATTCTG 1214
 DB 1096 GCTGCTCAACAGAGAGTGGGTGATCAACTGCGCCCTGACAGATGACATGACATTCG 1155
 QY 1215 TGGCTGTGACATGAATGCTCCCTGGGAGGTGCCAGATGCTGTGGAAATGCCGCTT 1274
 DB 1156 CCGGAGAGATTTCAACAGGCTCTGGGGGACAGATGACCATTAAGAGAGAGCGCCCTGT 1215
 QY 1275 CACGAGAGACAGGAGCCGACATGACGTCGTATGACATATATGTCTAACAGAACCATCTCT 1334
 DB 1216 GCTGGAACAAGATGATGCTGTGACCGCGTGGCTGTATGACTATGCGGGCGCACTGT 1275
 QY 1335 GGCCTTTGTGGGACCAAAAGTGGCAAGCTGAAAGAGATCCGGGTGATGAGACCA---- 1390
 DB 1276 GGTATTTGCGCGGACCGGAAAGTGGCGCATCCGCAAGATTTCTGTGAGCTCTCAACCC 1335
 QY 1391 -----GGGCAAGCCCTCCAGTATGAGACGTTGAGTGTGAGACCCCGCCAGTCT 1445
 DB 1336 CGTGGCGGGCTGTGCTGAGAGAGCTGTGAGAGAGCTGTGAGAGAGAGAGGAGGCCCATCT 1395
 QY 1446 CCGGATATGAGCTTCTTCAAGAGACAGAGACATCTTATCATGTCAAGAGAGAGAGCT 1505
 DB 1396 GCGAGACCTGTCTTCAAGCCCAACCAACAGTACCTTATGACATGACATGAGAGAGAGT 1455
 QY 1506 CACGAGTCCGCTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTG 1565
 DB 1456 GACCGGGGTGCTGTGAGAGAGCTGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTG 1515
 QY 1566 AGGAGACCCCACTGTGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1625
 DB 1516 ACGGAGACCCCACTGTGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1575
 QY 1626 TGAAGGTTCAAGAGAGCCCGGAGGTTGCTGTGAGAGTGAAGAGTGTGCTGCGCTGAC 1685
 DB 1576 TGAAGGAG 1635
 QY 1686 GATCATCCCAACATATCTCCGCTGTCAAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAG 1745
 DB 1636 TGTGACAGCCCGCAATGT 1695
 QY 1746 CAATGTCCCGAGAGTGTCACTGTGCGGACACTGTGACACTGTGAGAGAGAGAGAGAG 1805
 DB 1696 GAACTGTCTGACCTTCAAGTGTGAGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1755
 QY 1806 TGGGCTGTGTGTGAGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTG 1865
 DB 1756 GAGGCTGTGTGAGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGT 1815
 QY 1866 GATCATCAAGAGATGAGAGACCAACATGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAG 1925
 DB 1816 CATCAAGCGGGGCAAGAGAGACAGCGGGTGTGAACTCTTCAATTAAGTCAAGAGAGAC 1875

1926 CGGCATGACCTTCCAGACACACCTTGTCTTCTTACATTTGAGCGCTCCAACTTCGTG 1985
1876 AGGGAAGAAGTTTGCCTGTGTGACCTTGTCTTCTTACATGACGCGCTCCACCAAGTCTCTG 1935
1986 CTTGTCTCTGCTGGAAGTCCATACCGCTGCACCTGTGTAAATACCGGGAATGTCTGAC 2045
1936 CTTGTCTCTGTCAACGCTCTCTTCCCTGCTGCTGTGCAAAATACCGCACTGTGAC 1995
2046 CCATGACCCCAAGACCTCTCTCTTCCAGGAAGCCGAGTGAAGCTGCCAGGACTGCC 2105
1996 ACACACTGTGCTGCTGCTGCTCTTCTGTGAGGCGCTGTCAACCTGTCTGAGGACTGCC 2055
2106 CCACTGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2165
2056 ACAGATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2115
2166 GAAGGCAAGACCTCTCTCTTCCAGGAGCCGAGTGAAGCTGCCAGGACTGCC 2225
2116 GAGGCAAGACCTCTCTCTTCCAGGAGCCGAGTGAAGCTGCCAGGACTGCC 2175
2226 CATTGAGGCAAGACCTCTCTTCCAGGAGCCGAGTGAAGCTGCCAGGACTGCC 2285
2176 CATTGAGGCAAGACCTCTCTTCCAGGAGCCGAGTGAAGCTGCCAGGACTGCC 2235
2286 CCAAGAACCTCTTATTTCTTATGAAAGGATGGAATCAACACCTGCCGCTGAGTTC 2345
2236 CCAAGAACCTCTTATTTCTTATGAAAGGATGGAATCAACACCTGCCGCTGAGTTC 2295
2346 AGTCTGTGATGAGGCACTTCAACATTTGACCACTGAGTGAAGTGAAGTGAAGTGAAGT 2405
2236 AGTCTGTGATGAGGCACTTCAACATTTGACCACTGAGTGAAGTGAAGTGAAGTGAAGT 2355
2406 CAACTGTGAGGCACTGCTGAGGAGTGTGAGGCTGTGCTGAGTGAAGTGAAGTGAAGT 2465
2356 CAACTGTGAGGCACTGCTGAGGAGTGTGAGGCTGTGCTGAGTGAAGTGAAGTGAAGT 2415
2466 ATGTGCTGTGAGGCACTGCTGAGGAGTGTGAGGCTGTGCTGAGTGAAGTGAAGTGAAGT 2522
2416 GTGCGGATGTGTGCTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2475
2522 GAGGAGCCAGTGTGCTGAGGAGTGTGAGGCTGTGCTGAGTGAAGTGAAGTGAAGT 2582
2476 ACCTGATGCTGAGTGTGAGGAGTGTGAGGCTGTGCTGAGTGAAGTGAAGTGAAGT 2535
2583 AGAGATTAATCCGAGTGAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 2642
2536 CAACTGTGCTGCTGAGGAGTGTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 2595
2643 GAACTGTGCTGAGGAGTGTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 2702
2596 GAACTGTGCTGAGGAGTGTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 2655
2703 CAGCCTTTAGTGAATGTTACATCCCTGACAGACAGATGCTGTGTGAGTGTGAGTGTGAG 2762
2656 CAGCCTTTAGTGAATGTTACATCCCTGACAGACAGATGCTGTGTGTGAGTGTGAGTGT 2715
2763 CAAAGCC--AGCAGAGTGAAGGAGTGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGT 2819
2716 CAGCCTTTAGTGAATGTTACATCCCTGACAGACAGATGCTGTGTGTGAGTGTGAGTGT 2775
2820 ATTATGAGGAGGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGT 2879
2776 CTACGCGGCTCTTCACTTAAGGCTTCACTTCTGTGAGACCACTTCTTACGCTGAG 2835
2880 GCCCAGCAGGAGGAGGAGTGTGTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 2939
2836 CCGCTCCGAGGAGGAGTGTGTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 2895
2940 TGCAGGAGGAGGAGTGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGT 2999
2896 CGCAGGAGGAGTGTGTGAGTGTGTGAGTGTGTGAGTGTGTGAGTGTGTGAGTGTGTGAGTGT 2955

3000 TCCATCTTACATGTCTCTGACACACCATCTCTGAGTGAAGTGTGATGAAGTGTCTC 3059
2956 CTCCGTGAGATTCGGGTCTGACACACCCCGGAGAGCCCTGAGCAGCCTCCCATAT 3015
3060 GGTGAGGAGTGAACAGGAGCAGATTC--ACGAGACCTGTCTTTAGTATGTGAAGA 3116
3016 CATCAACATCAACGCGCCGACGCTTCAACACCTGTGAGTGAAGTGAAGTGAAGTGAAGT 3075
3117 CCGCAGCATCTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGT 3176
3076 CCGCAGCATCTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGT 3135
3177 CCGTATGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 3236
3136 GGTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 3195
3237 GAAAGGACATCAATATGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGT 3296
3196 CATTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 3255
3297 GCGCCTGCTGTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 3356
3256 GTTGTGTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 3315
3357 CTTGATCTGTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 3416
3316 CTTGATCTGTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 3375
3417 TCCCAACCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 3476
3376 CCGTACCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 3435
3477 GCGCATCATCTTAAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 3536
3436 CCGCATCATCTTAAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 3492
3537 GAACTACATGTGTGTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 3596
3493 CAACTACATGTGTGTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 3552
3597 GCTGTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 3656
3553 GCTGTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 3612
3657 GAACTACATGTGTGTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 3716
3613 CAACTACATGTGTGTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 3672
3717 CCGTACATGTGTGTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 3776
3673 TGTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 3732
3777 CTATTAACGCAATGCCGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 3836
3733 CTACAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 3792
3837 CAACTGTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 3896
3793 CAACTGTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 3852
3897 CATTGATGAGTGAACATGAGTGAACATGAGTGAACATGAGTGAACATGAGTGAACATGAGTGAAC 3956
3853 CATTGATGAGTGAACATGAGTGAACATGAGTGAACATGAGTGAACATGAGTGAACATGAGTGAAC 3912
3913 ATTATGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 3972
3957 TTACACATGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 4016
4017 GGTGAG 4076
3973 GGTGAG 4020
4077 CAAACAAAGTGTGTCTGTCTCTTCACTCCGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 4136

Db	4421	GACCAAGAGCACTTCCGTGCTGACCTTCATCGCAGCGCTGAGGACAGCGCAAGTCTTC	4080
Qy	4137	CATCCGCGACCGTGGCAACGTGGCTCACTCATATGACCGTGTGCAAGACCAAGCTGGA	4196
Db	4081	CATCGCGACCGCGGGAAATGTGGCTTCGCTCATATGACGGCCCTGCAAGGCGAGATGGA	4140
Qy	4197	GTAAGCCACTATGTGCTGAAGCAGCTGTGGCCGACCTCAATTGACAAAGAACTGAGAG	4256
Db	4141	ATAGCCCAAGGCGGTGCTCAAGCAGCTCTTCGACCTCATGCAAGAAAGAACTGAGAG	4200
Qy	4257	CAAGAACCAACCTTAAGTGTGCTGCTCAAGAGAGATGAGTCATGTGGCTGAAGATGTGAC	4316
Db	4201	CAAGAACCAACCCAAAGCTGTACTGTGCGCGAGCTGAGTGTGGTGAAGAAAGATGTAC	4260
Qy	4317	CAATTGGTTTACTTTCCTCTCTCTCAAAATTCCTCAAGAGAGTGTGTGGAGGCCCTCTT	4376
Db	4261	TAACTGGTTACCTTCCTCTTGTAAATGTTCTTCAAGAGAGTGTGTGGAGCGGCTGTT	4320
Qy	4377	CTCCCTGTTCTGTGCATTCAGACAGACAGATGGAAGAGGGCCCCATTTGACGCATCAAGG	4436
Db	4321	CATGTGTACTGCGGCATCAAGCAAGCAATGAGAGGGCCCCATTTGACGCATCAAGG	4380
Qy	4437	CGAGGCCGCTACTCTCTTGAGCGAGAGCAAACTCATCCGCCAGAGATTTGACTCAAAAC	4496
Db	4381	TGAGGCAAGCTACTCTCTGAGTGAAGACAAGTCATCCGCAGAGATTTGACTCAAAAC	4440
Qy	4497	CCTGTCTTGAGTGTGTGACGCCCAAGACAAATGCCAAAGCCCCGAGTCCCAATPAAGAT	4556
Db	4441	ACTGACCTTGAACTGTGTGAACCTTGAAATGAGATGACCTTGAGTGGCGGTGAAGG	4500
Qy	4557	CCTCACTGTGACACCATCACTCAGGTCAAGAGAGAAATTCGTGATGCCATCTTCAAGA	4616
Db	4501	GCTGAGCTGTACACGGTCAACCGGCGCAAGAGAAAGTGTGTGACGCTGCTCAAGG	4560
Qy	4617	TGTGCTTGTCTCCACCGGCCCCAAGCTGCAATTTGATCTGAGTGTGGCAAGAGAG	4676
Db	4561	CGTGCCCTACTCCCAAGCGGCCCAAGGCGCGAGACTGGAAGTGTGGTGGCCCAAGGCG	4620
Qy	4677	TGGGCAAGATGATCTTTGAGAGATGAAGATCAACCAAGATTTGAATTGGA	4736
Db	4621	CATGGCGCGCATCATCTGCAAGGACGAGAGAGTCAACCAAGATTTGAAGATTTGGA	4680
Qy	4737	GCGACTGAACACATGCGCCCATACAGGTGCAATGTTCCGTGTGGCAATTAGTTC	4796
Db	4681	GAGGTGAACACATGCGCTCATCTCAAGATGACACAGGATCTTGTGTGGCACTGTGTCC	4740
Qy	4797	CAAGCAGGTGACAGCTCTTAAACGAGTGAACAACTCCACGCTTCCAGAGCTTCAGCAG	4856
Db	4741	CAAGCAGAGTTCGCTCAACATCTCAACTCTTCCACTTCAAGAG---TCCCTAG	4792
Qy	4857	TAAATATTAATAACATGATCCGATACAGGAGAGCCCCGACGCTCCGCTCAAGCAAC	4916
Db	4798	CAGATACAGAGCATGTGCGACGAGCGACAGCCCCCAAGCTTCGCGCTGCGCACCC	4852
Qy	4917	TATGATCACTCTGACTTGAGAGTGTGAGTCAAGATGTGTGACACTTATGAAGAACAGCA	4976
Db	4858	CATGATCAAGCCCGACCTGAGAGCGGACCAAGCTGTGTGACCTGTGAGAAACACGA	4912
Qy	4977	GCAAGGAACCAAG	5036
Db	4918	CCACTGGAACAGCGGTGAGAGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG	4972
Qy	5037	CGACCTCTGAGCACTAAGAGGACACTGACAGAAAGTTTGTGATGACCTCTTGAAGCAT	5096
Db	4978	ACGCTACTAGGCAACAAAGGACACTGAGAAAGTTTGTGAGAGAGAGAGAGAGAGAG	5032
Qy	5097	CTTGAGCAAGGACACAGGTGTGCGCTGCGCTGCGCAATCAAGTACATGTTGAAT	5156
Db	5038	CTTGAGCAAGGACACAGGAGGTGAGCTGCGCTGCGCAATCAAGTACATGTTGAAT	5092
Qy	5157	CCTGATGAGCAGGCTGATTAACATGCAATTCATGACCCGACGTCGCAATCTGGA	5216

Db	5098	CCTGGAATGAGAGGGCCGACAAAGCACACCAAGTTCACGAATGCAGCATTGTGGCCACACTTGAAA	5157
Oy	5217	GAGAATTTGCCCTGCCCCCGTAGGTTTGTG3GATCAACATCATTAAGAACCCGAGTTTGTGT	5276
Db	5158	GAGACACTGCTGCTCCCTTGGCTTCTGTGGTGAAGCTATCAAGAACCCACAGTTTGTGT	5217
Oy	5277	TGACATTCATTAAGAACGACATCAAGACGCTGCTCTCTGTGTGTG	5322
Db	5218	CGACATTCACAAAGACAGATCAAGACGCTGCTGTGTGTGTGTG	5263

RESULT 13			
LOCUS	AKI22289	6893 bp	mRNA linear ROD 15-MAR-2003
DEFINITION	Mus musculus mRNA for mKIAA0463 protein.		
ACCESSION	AKI22289		
VERSION	AKI22289.1		
KEYWORDS	GI:28972234		
SOURCE	FII CDNA.		
ORGANISM	Mus musculus (house mouse)		
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus;		
AUTHORS	Okazaki,N., Kikuno,R., Ohara,R., Imamoto,S., Aizawa,H., Ynasa,S., Nakajima,D., Nagase,T., Ohara,O. and Koga,H.		
TITLE	Prediction of the coding sequences of mouse homologues of KIAA gene: II. The complete nucleotide sequences of 400 mouse KIAA-homologue cDNAs identified by screening of terminal sequences of cDNA clones randomly sampled from size-fractionated libraries		
JOURNAL	DNA Res. 10, 35-48 (2003)		
REFERENCE	2 (bases 1 to 6893)		
AUTHORS	Okazaki,N., Kikuno,R., Nagase,T., Ohara,O. and Koga,H.		
TITLE	Submitted (07-FEB-2003) Hisashi Koga, Kazusa DNA Research Institute, Laboratory for Genome Informatics: 2-6-7 Kanazu-kamatari, Kisarazu, Chiba 297-0818, Japan		
JOURNAL	(E-mail:muse@kazusa.or.jp, Tel:81-438-52-3915, Fax:81-438-52-3918) The CREATS program supported by Japan science and technology corporation; cDNA full insert sequencing; Kazusa DNA Research institute; cDNA library construction, clone selection and 5'- & 3'-end one pass sequencing.		
COMMENT	Location/Organifiers		
FEATURES	1..6893		
SOURCE	/organism="Mus musculus"		
	/mol_type="mRNA"		
	/db_xref="taxon:10090"		
	/cdate="mdg07539"		
	/tissue_type="brain"		
	/dev_stage="adult"		
	/note="vector:modified pBC SK+."		
	1..6893		
	/gene="mKIAA0463"		
	<1..4592		
	/gene="mKIAA0463"		
	/note="CDS is predicted by in silico analysis. Start codon is not identified."		
	/codon_start=3		
	/evidence=not experimental		
	/product="mKIAA0463 protein"		
	/protein_id="BAG5571.1"		
	/db_xref="GI:28972235"		
	/translation="QIKERLQSCFHGNGNLEMLLKDKVQCITAPVIPDNPFGDLID/NOPGLSPVPGSLILTYTSRDRLSVASVTYNGSVFVGTSKRKLKKIRADPDPHGGVOYMWVSVPFGSGELILDMAPSINQLLYWMSQRVTPVESCEQTTCGECLSSGGDAPIHSEIGIACAFGNLTVEGVGSVCVICSIPKPQVPIPLDQDFGELIQRSXETGKIPTSTEFNFYCSAHOLICSNVSAFCHEWCXYNLCHTDPTTOSFOGRINVSKDCPOLVTEBELIIPVEGKPTTLKANRLPOPSORGYECVLISQAVHRYPALRFNSISVQCNSSGYOQGMDSINTLADVPAVMNGFIINDODLVHLHYKCAAGCESGLCIKADPKREPCMGSGMRCTLIHQHCPSSTPWIMDSHWKCNPNQITELITATGSPPEGTRVTHIVNGLDPSSETLAHHYQVAGVCTPIRGATTAQCVCMGHAVICTGSPVNRTRICECKPEFTNTSHQQYTINBSVLSLPTRGESGNTVTTIGHITLGASSVAAYIHGNQCYEFSRSMETICVSPSSNGLDPVSVSDVDARVDSIQFEIIDDPVRQRIE		

PEWITSHTPLITITGNLDVIOEPRVKNESVNCVTNVTITLCLAPSLTSD
YRGLDVERPDEPFLFNNVOLLINDTKFYENPFEILSPGILDQSPBPI
LKKNICPSPAGAKNTVMIGETPTCTVSEBTLCEPPTLQGHKVMYVAGV
LSPGVISLITLPAITSIAGSLLIIVILIAKRSREMLDLKLOQMD
NLESRAVBEKENFALQTDINELSDLRSGIPYDVTVMRYLPEIEDHPLRE
LEVQNGQOQHEKALKU/ROLINKVPLITITLDELGSFMRBGNVSLIMTGL
GRLEYATDVKQLSLIDNLENNKPKLRTISVAKMLTMMFALHKLKKEC
AGEPLMLYCAIKQOMEKPIDAITEBARSISEDLIQQOIEKTLINCVNPDEN
SPEIPVYLNCDTITVKEKIDAVKVMYVQSPRAVMDLEMRQIARVLDND
ITRIEDGMRKLTIMHVOYSRVVALVYKOTSSYNIPASISRTSISRDSFRY
TSPDLSRSPWMTPELBSGVKMLVKNHDKQKEDGRSKMVEIYLRILATK
GTOKRPTDILFETLPSVHSGALPIAICKMPDLDEQDHSIHDTDVHMKNSCL
PLRPWNVIKRPOFPIHKSITTDACLSTVATPMDSSSTSHRIGDPSNKLLYA
KDIPLSYNWEYIADIAKUPALSDOMNAYLAEQSLHATEFNMLSALNEIYSYVK
YSEBILGALSDQDRQRRLAYVEHLINMSTES"

ORIGIN

Query Match: 42.0%; Score 2387.8; DB 10; Length 6893;

Best Local Similarity 70.6%; Pred. No. 0; Mismatches 1342; Indels 12; Gaps 4;

Db	1099	ATTAGAGCGCGTGCAGTCTTTGTACCGGGCGAGGGCA	CGCTGACCTGCGCTGAGCTC	1158
Qy	6	ATCAAGAGCGGTGCAGTCTCTGCTACATGAGAGGCA	CTTGAGAGCTCAACTGAGCTG	65
Db	1159	AAGGTGAAGACATCCCTCGAGAGAGTCCGCTCTTA	ACCATGACATTACTTCTGAGG	1218
Qy	66	CTGGGAAAGATGAGAGTGCAGAGTGCACCAAGCG	CGCTGTCCCAATCGATTA	125
Db	1219	CTGACATGATGATGCTCCCTGGAGTGCAGACATG	TGCGTGAATTCCTGCTTCA	1278
Qy	126	CTGACATGATGATGCTCCCTGGAGTGCAGACATG	TGCGTGAATTCCTGCTTCA	185
Db	1279	GAGACAGGAGACCGGATGACGTCTGTCAATGCA	ATGCTCAAGAACCACTCTGAGG	1338
Qy	186	ACGACAGGAGACCGGATGACGTCTGTCAATGCA	ATGCTCAAGAACCACTCTGAGG	245
Db	1339	TTTGTGGGACCAAAAGTGCAGAGTGCAGACATG	CCGAGTGCAGAGGAGGCAAC	1398
Qy	246	TTTGTGGGACCAAAAGTGCAGAGTGCAGACATG	CCGAGTGCAGAGGAGGCAAC	305
Db	1399	GCCCTCCAGTATGAGACGCTGACAGGTGT---	GACCCCGGCGGATGCTCCGAGATG	1455
Qy	306	GCGGTCCAGTATGAGTGTCTGTGTTCAAAAGT	GAGACCCCAATCTCCGAGATG	365
Db	1456	GCCCTCCAGTATGAGACGCTGACAGGTGT---	GACCCCGGCGGATGCTCCGAGATG	1515
Qy	366	GCCCTCCAGTATGAGTGTCTGTGTTCAAAAGT	GAGACCCCAATCTCCGAGATG	425
Db	1516	CCTGTGAGTCTGTGTGATGATGAGTGCAGAGT	GCCTGAGTGCAGAGGAGGCT	1575
Qy	426	CCTGTGAGTCTGTGTGATGATGAGTGCAGAGT	GCCTGAGTGCAGAGGAGGCT	485
Db	1576	CAGTGTGCTGTGTGTGCTGCAACAGTGCAGAG	CCGAGGAGGCTGAGCGGTCC	1635
Qy	486	CAGTGTGCTGTGTGTGCTGCAACAGTGCAGAG	CCGAGGAGGCTGAGCGGTCC	545
Db	1636	AAGGAGCCCGGAGGTTTGTCTCGAGATGAAG	CAAGTGTGTCTCGAGTGCAG	1685
Qy	546	TGGGAGCAAAATGATTTGTCTGCGAGTATCAG	CCAGTGTGTCTCGAGTGCAG	605
Db	1686	AACATATCTCGCTGCTCAGTACAGAGTGCAG	AGTGTGTCTCGAGTGCAG	1755
Qy	606	AACATATCTCGCTGCTCAGTACAGAGTGCAG	AGTGTGTCTCGAGTGCAG	665
Db	1756	GAGCTGTCAAGTGCAGTGCAGCTTTGAGAG	CTGTCAAGAGATGAGTGTGCT	1815
Qy	666	AACCTCTCGAGAGTATGCTGTGTGCTTGG	AAATCGAGTGTGAGTGTGCT	725
Db	1816	GTGGGCAATCAATCAAGTGCAGTGCAGGCA	GAGTGTGAGTGTGAGTGTGCT	1875
Qy	726	TCTGGAGTCAAGTGCAGTGCAGTGCAGGCA	GAGTGTGAGTGTGAGTGTGCT	782

Qy	1876	GAGATGGGAGCCACCAATGTGTGACAGTTCAG	CTCAATCAAGAGGAGCGGATGAC	1935
Db	783	CTGATCAAGCTGTTGGCTTGAAGTGCAGTGT	GAATCAAGAGGAGGAGGAAATG	842
Qy	1936	TTGCCAGACACAGCTTTGTCTTCAATGCAAT	TGCAAGCTGCAATTCGTGCTGTCC	1995
Db	843	TTTGTGACGACCAAGATTCAGATTCATAGC	AGTGTCCCAACAGTGTGCTGTCC	902
Qy	1996	GTGAGAGTCCATACCGCTGCTCATGTGTAA	TACCGGAGATGCTGCAACCATGACCC	2055
Db	903	GTAAACAGCGCTTCCGTGCTCATGTGTGCA	AGTACCGTAACTCTGACATGACCCC	962
Qy	2056	AAGACTGCTCTTCCAGGAGGCGGAGTGAAG	CTGCGCCGAGACTGCCCCAGCTGTG	2115
Db	963	ACTACCTGTTCTTCCAGGAGGAGAGATCA	ATGTTTCAAGAGACTGTCCCCAGCTGTG	1022
Qy	2116	CGAGTGACAAAGTCTGTGTGCTCCGTGAG	GTATCAAGCTTATCAGCTGAAAGCCAG	2175
Db	1023	CCCAAGGAGGAGATTTGATCCAGTTGGG	AGATTAACCAATCACCTTAAAGCCCGA	1082
Qy	2176	AACCTCCCCAGCGCCAGTCTGGGAGCGT	GGTACGATGATCATCTCAACATTCAAGGC	2235
Db	1083	AACCTGCCCGAGCCCAATCTGGGAGCG	ATACAGTGTGTGCTCAGCATTCAGAGG	1142
Qy	2236	AGCGAGCGAGTGCCTGCTGCTTCAACAG	CTTCAGCTTCAAGCTTCAAGTGCACACC	2295
Db	1143	GCTGTCCACCGGCTTCCCTGCTGCTTCA	ACAGTTCAGATGTCAGTGCACAAACG	1202
Qy	2296	TCTTATTCCTATGAGAGGATGAGATCA	CAACCTGCGCGGAGTGTGACAGTGTG	2355
Db	1203	TGTACCAATGATGATGATGATGATGAT	GATGATGATGATGATGATGATGATG	1262
Qy	2356	AATGGCACTTCAACATTGACACACCACT	CAAGATTAAGTTTCACTTAAAGTGTGA	2415
Db	1263	AATGGCACTTCAATATGACACACCACT	CAAGATTAAGTTTCACTTAAAGTGTGA	1322
Qy	2416	GCTACGCTGAGAGTGGGCTGTGCTCAAG	GTGACCAAGCTTGCATGTGTGCTG	2475
Db	1323	GCTACGCTGAGAGTGGGCTGTGCTCAAG	GTGACCAAGCTTGCATGTGTGCTG	1382
Qy	2476	TGCCAGGCGCCAGGCGCAGTGAACCTG	CGCAGCATGCTGCTGCCAGAGACCTG	2535
Db	1383	TGCAGTGTGAGCGAGATGATACCTT	CCACAGCATGCTGCTGCCAGACCTG	1442
Qy	2536	CTGAGAGCTGTCTGTGCTCAAAAGTGC	CAAAACCCCGCATGACAGATTAATCCG	2595
Db	1443	CTTGAATGTGCAACAAATGATCAAGTTC	CAACCCCAATCAAGATTTTGA	1502
Qy	2596	GTGACAGGCGCCGAGGAGGAGGACCA	AGGTCAATTCGAGAGGAGAACCTG	2655
Db	1503	GTATGAGGACCACTGAGAGGAGGAC	CTGTGATCATTCATGAGGCTG	1562
Qy	2656	GAAATTCGCGATGCTTCCATGATCAAG	TGTGTGCTGAGTGTGAGTGTGAGT	2715
Db	1563	GAAATTCGCGATGCTTCCATGATGAGT	TGTGTGCTGAGTGTGAGTGTGAGT	1622
Qy	2716	GATGTTATCATCTCTGACAGACAGAT	TCGTGTGTGATGAGTGTGAGGAGG	2775
Db	1623	GAGGAATATCATGCTGTGAGAGATTC	GTGTGTGATGAGTGTGAGGAGG	1682
Qy	2776	CATGAGCTTGTGTGAGATCTGTGCTGT	GTGTGCTGTGCTGTGCTGTGCTGTG	2835
Db	1683	ACATCTGGGCGCTGTGCTGTGATGTG	GGAGATGCAAGGATTCATGACCAATG	1742
Qy	2836	TCAAGCTCTATCTTCTTCACTGAC	CTCTCTGATCTGATCTGATCTGAT	2895
Db	1743	CACCAAGATATCTTGTGTGATCTTCT	GTGTGCTGTGCTGTGCTGTGCTGTG	1802
Qy	2896	ATGTCGAGAGGACCAATGATGACAT	CAACAGGACCAACTGAAAGCGGAG	2955
Db	1803	GAGTCAAGAGGATCAATGATGATCA	AGGACCACTTACCTTGTGTGTGAG	1862
Qy	2956	GTGGTATGTGTGAAAGAGCGCTGT	CTTCCACAGGCGATTCATCTTATG	3015

Db 1863 GCACGTGACCTGGGCAATCAGACCTGTGAATTTCTATGSGAGCAATGATGATTTGTA 1922
 QY 3016 TG---CAACACACATCCTCAGATGAGGTGCTAGATGAAAGTGTGCTGACAGTGGAC 3072
 Db 1923 TGTGTTTCAACCCCATCATCAATGACTAGAACCAAGTCCCTGCTCTCGTAGGTGTGAC 1982
 QY 3073 AGGGGCAAGATCCACAGAGACCTGTGCTTTCAGATGAGTAAGAACCCACCATGTGCGG 3132
 Db 1983 AGAGCCCGGATGATGACAGTGTGCAATTTGAGTATTAATGACCCAGGGTCCACGT 2042
 QY 3133 ATTGACCAAGATGAGCATTTGCTAGTGAACAACACCAATCGCCGATGAGGAGACCCAC 3192
 Db 2043 ATTGACCAAGATGAGTATCACTAGTGGGACACACCCCTAACCATCAAGGCTTCAAC 2102
 QY 3193 CTGACCTCATACAGAACCCCAAGATCCGTGACCAAGATGAGGAGGAGACATCAAT 3252
 Db 2103 TTGATGTCAATTCAGAGACCCAGGATCCGAGTCAAAATTTAATGCAAGAAATCTGCAAT 2162
 QY 3253 ATCTGTGAGTTTGAAGCTACTAGATGACCTGTGAGGCGCCGCGCTCGTCTGGGT 3312
 Db 2163 GTATGCAAGTGTAAACAAACACCCCTCACTGTGTGGACCCCTCTGACCAAGTAC 2222
 QY 3313 CCTGACCAACAGTCAAGCTTCAACGAGAGCCCAAGAGATTGGCTTCAATCTGACAC 3372
 Db 2223 TACCGCCCAAGGTCTGAGCACTGTGGAACGGCCAGATGATTGATTTCTTTAACAAT 2282
 QY 3373 GTCCAGTCCCTGCTATCTCAACAGACCAACTTCACTTATCCCAACCCGCTTTT 3432
 Db 2283 GTTCAATCTTACTATCTATTAAGACACCAAGTTCACTTACACCCCAACCTTT 2342
 QY 3433 GAGGCTTTGGTCCCTCAGAAATCTGTGAGTCAAGCTGACGCCCATCATCTTAAAG 3492
 Db 2343 GAACGTCTCAGGCCCACTGGAATCTTGGATCAGAAAGCCAGGCTCAACCATATCTTAAG 2402
 QY 3493 GCGAAGAACCTGATCCCGCTGTGGCTGGGGCAACGTGAAGCTGAATCACTGTGCTG 3552
 Db 2403 GCGAATAATCTGTCTCTCTGCTGCTGAGG--GCGCAATCTCACTACACGTAATG 2459
 QY 3553 GTTGGGAGAAAGCCGTGACCGGTGACCGTGTCAATGTCTCAGCTGTGAGTCCCC 3612
 Db 2460 ATTGAGAGACACTTTGTAAGTCACTGTGTCTGAGACACAGCTGCTTTGGAACCTCC 2519
 QY 3613 AACCTCATGTGAGGACCAAAATGATGACCGCTGTGCTGTGAGTACTCCCGGG 3672
 Db 2520 AACCTCAAGGACCAAAAGTCAATGTGTCAGTGGGCGAGATGATGTCTACCTGAC 2579
 QY 3673 ATGTGTACATGCCCCGACAGCCGCTCAGCTGCCCCCATCTGTCAGATGCGAGTG 3732
 Db 2580 TCCGTGAGGTATCTCCGACAGCTGTGTAACCTGCGACGATCATGACATGCGAGCT 2639
 QY 3733 GCTGGCGGCTCTCTATCATTTTTCATCGTGGCCGTGTCTATTTGCTTAAACGCAATCC 3792
 Db 2640 GGTGGAAGCTCTCTTATCATCTGTCAATTTGCTTCATCGCTTAAACGCGAAGCT 2659
 QY 3793 GCGGAAGTGAACCTCAGCTGTAAGCGGCTGAGATGAGTGAACAACCTGAGTCCGT 3852
 Db 2700 AAGGAATATGACTCATCTCAAGGGGTCCAAATGCAATGAGACACTGGAATCCAGG 2759
 QY 3853 GTGGCCCTGTGATGCAAGAGACCTTTGCGAGCTGACAGCGGACATCATGAGTGAAC 3912
 Db 2760 GTGGCACTGAGTGAAGAGACCTTTGCGAGCTTCAAGACAGATCAATGAGCTAAC 2819
 QY 3913 AATGACCTGATGAGAGCGGAGATCGGTTCTGTGACTATAGAACTTAAACCATGCGGTTG 3972
 Db 2820 AATGATTTGATGATCAGAAATCCCTTACCTGAGACTACGTTAATGCACTGAGATC 2879
 QY 3973 CTGTTCACAGAAATTAAGACCACTCTGTCTCCGGAACCTTGAAGTCCCGGATCCGG 4032
 Db 2880 CTGTTCACAGGCAATTAAGACCACTCTGTCTCCGGAACCTTGAAGTCAAGGAAATGGA 2939
 QY 4033 CAGGAGCTGTGAGAAAGGCTTGAAGCTTTGCGCCAGCTCATCAACAAGGTGTTC 4092

Db 2940 CAGCAGACGTGAGAAAGCCCTGAACCTTTCGCCAGCTTATCAACAAGGTGTTC 2999
 QY 4093 CTGCTGTCTCTCAATCCGCAAGCTTGAAGTCCAGGCTACTCTTCATGCGGACGTGAC 4152
 Db 3000 TTGCTGACCTTCTATCTGACACTGGAACCTCAAGCCAGCTTCTCATGAGACCGTGG 3059
 QY 4153 AACGTGCTCACTCATATGACCGTGTGCAAGAACAGCTGAGTACGCACTGATGTG 4212
 Db 3060 AACGTGCTCTCTCATATGACAGGCTTCAAGGCTGCTCAATATATGCACTGATGTG 3119
 QY 4213 CTGAAGCACTGCTGCGGACCTCATTTGACAAAGACCTGAGAGAGAAACACCTTAAG 4272
 Db 3120 CTCAAGCACTGCTGCTGACCTCATTTGACAAAGCTGAGAAACAAACACCTCAAG 3179
 QY 4273 CTGCTGTGAGAGAGCTGAGTCACTGAGCTGAGAAAGTCTGACCAATTTGTTACTTTC 4332
 Db 3180 CTGCTTCTCCGAGAGACTGAGTCTGTGGCCGAGAAAGTCTGACTTAACTGTTCTT 3239
 QY 4333 CTCTCTTACAGTTCTCTCAAGAGTGTGTGGGAGCCCTTCTCTCTGTCTGTGCG 4392
 Db 3240 CTTCATACAGATTTCTTGAAGAGTGTGTGGGAAACCACTTTCATGCTATATGTGCA 3299
 QY 4393 ATCAAGCAGCAGATGAGAAAGGCCCATTGAAGCATCAAGGCGAGGCCGCTACTCC 4452
 Db 3300 ATCAAGCAGCAGATGAGAAAGGCCCATTTGACCTTATCTGAGAGGCCGAGATCTCC 3359
 QY 4453 TTGAGGAGAGCAAGTCACTCCGACAGAGATTGACTACAAACCTTGTCTGTAGCTGT 4512
 Db 3360 CTGAGTGAAGAGAGTCACTCCGAGAGAGTCAAGTATTAACCTGATCTGTGAAGT 3419
 QY 4513 GTCAAGCCAGAAAGGCCAAGCCCGAGTCCAGTAAAGTCTCACTCACTGACACC 4572
 Db 3420 GTCAAGCCAGAAAGGCCAAGCCCGAGATCCAGTAAAGTCTCACTGACACC 3479
 QY 4573 ATCACTCAGTCAAGAGAGATTTGTGATGCTATCTTCAAGATGTGCTTGTCTCAC 4632
 Db 3480 ATCACTCAGTCAAGAGAGATCTCGATGCGGATATTAAGATGCTCCCTACTCCAG 3539
 QY 4633 CGGCCCAAGCTGCAATGTGATCTGAGGTGGCAAAAGAGGGGCAAGATGATC 4692
 Db 3540 CGGCCAAGCTGTGACATGATCTGAGGTGGCCCAAGCCGATGCCAGATGTG 3599
 QY 4693 TTGCAAGTGAAGATCAACCAACCAAGATTGAAATGATTGAAGCACTGAACACTG 4752
 Db 3600 TTGCAAGTGAAGATCAACCAACCAAGATTGAAATGATTGAAGCACTGAACACTG 3659
 QY 4753 GCCCACTACAGGTGCAATGTTCCGTGTGCACTTGTGCAAGAGGTGACAGCC 4812
 Db 3660 ATGATTTACAGGTGCAACAGATGCTGTGAGCTGTGCTTCAAGCACTGTCTCC 3719
 QY 4813 TATACGCAAGTGAACAATCCACCGTCTCAGAGCTCAGCAAGTAAATATGAAACATG 4872
 Db 3720 TACAAATCTCTGCTCTGCGACAGATCTCTCGGATTCATTAAGAGATGACTCTTCC 3779
 QY 4873 ATCCGATACCGGGCAGCCCGGACAGCTCTCGTCAAGCACTTATGATCACTCTGAC 4932
 Db 3780 TTCAAGTACCAAGGACCCCAAGACAGCTCTCGGTCTCCGATCCCAATGATCCCAAG 3839
 QY 4933 CTGAGAGTGAAGTCAAAATGTGAGCACTAGTGAAGAACCAAGAGCAGAGCAGAAAG 4992
 Db 3840 TTGAGAGCGGTGTCAAGGTTTGGCACTGTGTGAAGAAATCAATGACATGTGACGAAG 3899
 QY 4993 GAGGGGACCGGGGAGACAGATGTGTGTGAATTTACTGACCCGACTCTGTGCACT 5052
 Db 3900 GAGGGTACCGGGGACAGAAATGTGTGTGAATTTACTGACCCGCTTCTAGACACC 3959
 QY 5053 AAGGGCACTGAGAGATTTGTGATGAACCTTTGAGACATCTTACAGCAGGACAC 5112
 Db 3960 AAGGGCACTGAGAGATTTGTGATGAACCTTTTGAACCTTTTACAGCAGGACAC 4019
 QY 5113 CGTGTCTGCTCCCTGCGCCCTGACATCAAGTATCATGATTTGATCTCTGATGAGCAGCT 5172
 Db 4020 CGGGTGTGTCTCTCCCTTACGATCAAGTATCATGATTTGATCTCTGATGAGCAGGCA 4079

Db	53	CAAGTTGGGGAGGACCTCCACACCGAAGAAACATTACCTTTCTAAGTGTCAATGAGCTCG	589
OY	501	CTGAGCTTTTGGAGGATATGTCTCTCTTAAGACACCTGGATGACAGCTGTTCATGCGAC	560
Db	590	GACCAATGTCTGGTGTCATCATTTAGAGTTCCAAATGGCCAAATAATPAACTCTTTGTGGAGC	649
OY	561	GGCAGTGGATGGGAAGCCCGAGTATTTTCCACCATCTCCAGCCGGAAACTGACACAGAA	620
Db	650	TCCATATGATGAAATAATCTGAATATTTTCCACACCTCTCCAGCCGCAACTACTGGGCA	709
OY	621	CTCTGAGGCGGATGGGATGTGTGGCTGACCTTCATGAGAGTGTGGTGGCCCTGCATGAT	680
Db	710	TGAAGAAATGCTGAATATGTTGGCTTTGTCTATACGGATGAATTTGTGTCTTCACAACT	769
OY	681	TAAATTCCTTGGGACACCTTCACCATCATCCTGACTTGTATATGTAATGTGTATGG	740
Db	770	CAAGATCCCTCCAGATACTTGTCAAAATTTCCCTCAATTTGATATCTAATCAATGTGAAG	829
OY	741	TTTTAGAGAGGGAATTTGTCTACTTTTATACCTCCAACTGAGATGAGTGTCTCCACC	800
Db	830	TTTTCAGAGTGAAGCACTTTGTGTACTTACTTACCTCTCAGCTGACACCAATTAATC	889
OY	801	AGGCTCCACCAACCAAGAGCAGGTGTATATCCAACTGTAAGCTTTGCAAGAGAA	860
Db	890	ACCTGACTCAACAGGGGACAGTTTTTTTAACTCCAAATGTGCTCCGCTTTGTGTAGCA	949
OY	861	CACAGCTTCACACTCCTATGTAGAGTGGCATTTGGCTGTGAGCGAGTGGGTGGAGTA	920
Db	950	TCCCAAGTTCTACTCTTACGTAGATTTCCATTTGGTGTGATGAAGATGGGGTGAATA	1009
OY	921	CGGCTGTGCAAGCTGCTTACCTGTGCCAAAGCGGGGCGGTGCTTTGGACGACCTTTGG	980
Db	1010	TAGGCTGATACAGGATGCTTATCTAGACAAACCAAGGAAACGCTTTGGCTTAAGAGCTGG	1069
OY	981	AGTCATCCAGATGATGACCTGTCTTACCGCTTTTCCAAAGGGCCAGAGGGGAAAT	1040
Db	1070	CATCTGTGAAGGGAAGACATTTCTTTCAACGTCTTTCACAAAGACAAAGAACCGGAT	1129
OY	1041	GAAATCCTGGATGAGTGGGCTCGCTGTGCATCTTTCATCTTGAAGCAGATTAATGACCGAT	1100
Db	1130	CAAAACCCCCCAAGGATCTGTCTGTGGCTCTTCACTTGAAGAAATCAAGGATTAAT	1189
OY	1101	TAAAGAGCGGCTGCACTCTTTTACCGGGGGAAGGACACCTGACCTGGCTGGCTCAA	1160
Db	1190	TAAAGACCGATTCAGTCTTGTCTTAAGAGGGAATGGGAAATCTCTCTGCTTGGCTCTT	1249
OY	1161	GGTGAAGACATCCCTGTCAGAGTGGCTCTTAAACATTAAGATGACATTACTTGTGGCT	1220
Db	1250	AAACAAGAGCTGGGATGATCAATTCACCTCTCAATTAAGACAAATTTCTGTGGGA	1309
OY	1221	GGAATGAATGCTCCCTGGAGTGTCCGACATGATGCGTGAATTCGCCGTCTTACGGA	1280
Db	1310	GGAATTTAACACACCGCTTTGGGGGGAACATGACATAGAAAGTACACGTGTGTCTGGA	1369
OY	1281	GGAAGGAGCCGATGACGTGTGTATGTCATATGTCATATGTCTAACAGAACATCTCTGGCCTT	1340
Db	1370	TAAAGAAGAGGGAATACATCAGTAGCGGCTTAATATATGCGGGGACACATGTGTGT	1429
OY	1341	TGTGGGCAACCAAAATGGCGAAGCTTAAGAAATCCGGGTGATGTGACCAAGGGGCAACGC	1400
Db	1430	TGCTGGAACCCCGAGTGGCCGTGTAAAAAATCTGTGTGAACCTGTCTGGCTCTGCTC	1489
OY	1401	CTTC-----CAATATGAACCGTGCAGTGTGTGAACCCCGGCCAATCTCTCCGGA	1451
Db	1490	TCACTTGGTGCAACATATGAGAAATGTGTGTGTCCATGAAGGCGAATGTATCTTGAAGGA	1549
OY	1462	TAAAGCTTCTCCAAAGAACAGACAACTATACATATCTCAGAGGAGAGCTCACAG	1511
Db	1550	CTTGTATCTGAGCCCGGACCGGCAATATATCTAAGCTATGACAGAAAAGAGATTATCTG	1609
OY	1512	AGTCCCTGTGGAGATCTTGTGTCTAGTATACAGAGCTGGCGGAGTGTCTTGGCTCAGGGA	1571
Db	1610	TGTCCCGGTGAAGTGTGTGAACAATATGAAGAGCTGTGACACCTGCTTGTGGCTCCAGGGA	1669

QY	1572	CCCCACCTGAGGCTGAGTGTGCTGAGCAACAAGCTGACCCGGAAAGAGCGGTGACGG	1631
Db	1670	CCCTCACTCGGCGCTGTGTGTGTGTCTCCCAACATGTGTTCAGAAAAGATTAATGCGACGG	1729
QY	1632	GTCCAAAGAGCCCCCGAGGTTTGCTCCGAGATGAAAGACAGTGTGTCCGGCTGACGGTCA	1691
Db	1730	AGCAGATGAGCTTCAACAGGTTCACTTCTGTGACACAGCCGTCAAGTGTGTCCACTCCACAGTCA	1789
QY	1692	TCCCAACAATATCTCCGCTCTGAGTCAACAAGTCTGTGCTCTGTGAGACGATCAATGT	1751
Db	1790	CCCCAAAATATCTCCGCTCACTGTGTCCAGGTTTCCATGGTTTACAAAGCTGTGATGT	1849
QY	1752	CCCGAGCTGTCACTGTGCGCTCAACTGACCTTTAGAGACCTGTCAAGATGATGGGCT	1811
Db	1850	CCCTGACCTGTCTGTCTGGGGTGAACCTGTCTCTTTAGAGATTTAATAAGATGGAAGAGCG	1909
QY	1812	GGTGTGGGCAATCAATCCAGATGCTACTCTCCCTCGACGCCAAGAGAGTGCCCGCGATCAT	1871
Db	1910	GATACTGATGGAAGATATCTACTGCACATCTCCCTCTGTCAAGAAATTTATCTCTATCAC	1969
QY	1872	CACAGAGATGAGGAGCCACCATGTCTGACAGCTTCAATCAATCAAGAGACCCGCAT	1931
Db	1970	AAGGGCCCATGGAACAAGAAGGTTGTAAACCTCTACTTAATCCAAAGAAATGTGGAA	2029
QY	1932	GACCTTGCCGACGACACAGCTTTGTCTCTCAATTGAGAGGTCCACAATTCGTCTCTGTC	1991
Db	2030	GAACTTTTGCAAGTGTGATTTTGTCTTCTCAACTGATAGTGTCCATCAGTATATCTCTC	2089
QY	1992	CTGCTGTGAGAGTCCATACGCGCTGCACTGTGTAAATACCGGCATGTCTGCATCCCATGA	2051
Db	2090	CTGTGTAAATGAGATCTTTCCCTCCGCACTGTGTGAATAATAGACACGTGTGCACCCCAA	2149
QY	2052	CCCCAAGACCTGTCTCTTCAGAGAAAGCCGAGTGAAGCTGCCCGAGACCTGCCCCACGT	2111
Db	2150	TGCAGCCGACTGTCTCTTTCAAGAGGSCCGAGTGAACATGTCCGAGGACTGTGCCCCACAT	2209
QY	2112	GCTGCGAGTGAACAAGATCTGTGTGCTGTGCTGTGAGAGTATCAAGCTATCAAGCTGAAGGC	2171
Db	2210	TCTCCCTCTCTCCGAGATCAACATCTCTGTGTGGGTGTGAAGCCCATCACTCACTGACTCG	2269
QY	2172	CAAAACCTCCCCAGAGCCCAAGTGTGGGACAGCTGTGACGAATGCAATCTCTCAACATTTCA	2231
Db	2270	CAAGAACCTCCCGAGGCCCAAGTCCGCTCAGCCGCACTTACGAGTATTTTCCACATCCC	2329
QY	2232	GGGCAAGAGAGCGAGTGTCCCGCTGCGCTTCAACAGCTCCAGGCTAACAGTGTCCAGAA	2291
Db	2330	CGGAGAGTGTACTGGGTCACTGTGTTTGTGCGTTCACAGTACACAGCATCCAGTGTCCAGAA	2389
QY	2292	CACCTCTTATTCATATGAAGGATGAGATCAACACCTGCCCCGTGAGTGTGACAGTGT	2351
Db	2390	CACCTCGATATCAATATGAAGCAATGATATCAGTGACTTACCTGTCAATTTATCTGTGT	2449
QY	2352	GTGAATAGGCGACTTCAACATTTGCAACCCAGCTCAACATTAATTAAGTTCACCTTCAAAAGTG	2411
Db	2450	CTGAATAGGCGACTTGTCTATGACCAACCTCAAAAATTCATCAACTCATCTGTCAAAAGTG	2509
QY	2412	TGAGAGCATGCGTGAAGCTGTGGGCTGTGTCTCAAGCTGAACCCAGACTTGTGCATGTGG	2471
Db	2510	CTGGGCCCTTAAGGAAAGCTGTGGGACTTTGTGCTTAAATCTGAACGCTGATTTTGAGTGTGG	2569
QY	2472	CTGTGTGCCAGGCCCCAGGCCAGTGTCAACCTGTGCGGACACATGCGGCTTCCAGAGAGACCA	2531
Db	2570	CTGTGTGCTCAGTGAAGAAAGTGTCACTGTGCGCCAAATTTGTCCACCTTGAAAAACC	2629
QY	2532	GTGTGCTGAGCTATCTGTGTGCCAAAGCAAGTGTCAAAACCCCGCATCAACAGATTAAT	2591
Db	2630	TTGATGTATCTCACTATCCGCAATATAGCTGTGCTGACTGACCCAAAGATATCAACAAAGCTGT	2689
QY	2592	CCCGGTACACAGCCCCCGGGAAGGGGACCAAGGTCACTATCCGAGGGGAGACCTGGG	2651
Db	2690	TTCAGAGACAGCCCCCAGGCGGTGTGAACCCGACTACATCAACAGAGAAATCTGGG	2749

QY	2655	CCGGAATTGGGACATCCGCTCCACATGTCGAAAGTTGGTGGGGTGAAGGAGCCCTT	2711
Db	2750	CTTGAAGTTCAGAGACATTCGCTTTGAGTGAAGGTTGGGACAGTCATGTGTCCCTGT	2809
QY	2712	AGTGAATGTTACATCCCTCGAGAAACAGATCGTGTGTGAGATGGGGAAGGCC--AAGC	2768
Db	2810	CGAGAGCGAGTATATCAGCGCAGAGCAGATCGTTTGTGAATTAAAGATGCCGGAAGAGC	2869
QY	2769	CACGACGATGACAGGCTTCGTAAGATCTGTGGGTGTGTGTCCGCTCAATTATGATGC	2828
Db	2870	CCGTGTCCAGAGGCTCCTAAGTGAAGTCTGTGTAAAGATCTGTCTCAAGACTTAAGAGCC	2929
QY	2829	CCGATCTCAAGCTCTATATCTTCATGACCTGACTCTCTCAGATCTGAAGCCAGCCG	2888
Db	2930	TATTTCTCAAAAGTCAATTCACCTTTGTGTACCTCTCTTAAACCGGATCAACTCTTCCG	2989
QY	2889	GGGGCCATATGTCGAGGAGACCCAAATGACATCAGAGCAGCAACTGATGCGGAG	2948
Db	2990	AGGTCTTTGTCTGTGGCACCCTGGATATCATATGAAGAAATTACTGAAGCTGGCAG	3049
QY	2949	CAACGTGTGTGTGATGTTTGGAAAGCAAGCCTGTCTTCCACAGCCGATCTTCATCTTA	3008
Db	3050	TGATGTGTTCTGTGGCTATGAGGAGACGTCCCTGTGATTTTCAATGCGAATGTCMAAGA	3109
QY	3009	CATTGTCTGACACACACATCCTCAGATGAGTCTAGATGAAGTGTCCGATCAGAT	3068
Db	3110	GATCCGCTGTAAAGTCTCCGACAGGACCAAGACCGGTAAGCCGAATTAACAATCTCAT	3169
QY	3069	GGACAGGCGCAAGAT--CCACAGAGACTGTCTTCAATATGTGAAGACCCACCAT	3125
Db	3170	CAACCCGCGACAGATGAACAATCTGAAGTCAATTAACAATTACACGAAGATCCACCGT	3229
QY	3126	CGTGCAGATTGAGCAGATGAGCATGTCAATGAGAAACACACCACATGCGGTATGGG	3185
Db	3230	CGAAAGATAGAACCAGAGTGAAGATGCGCAGTGGGGGAACCCCTCATATGTAACG	3289
QY	3186	GACCCACTGACCTCATACAAACCCCAATCCGTGCCAAGCATGAGGGAAGAGACGA	3245
Db	3290	GATGAATCTGGCACTATTAAAGAACCCAAATAGAAGCTTAATTAAGAGATGTAAGAA	3349
QY	3246	CATCAATATCTGTGAGGTTCTGAACGCACTAGATACCTGTCAAGCCGCCGCTCGC	3305
Db	3350	AGAAATTAAGTGAACCTTTAACAATGACCAACCATGTGTGTGGTGCACCCCTCGTTGA	3409
QY	3306	TCTGGGTCTGACACACAGTCAAGACTTAACGAAAGCCCGAAGATTTGGCTTCATCT	3365
Db	3410	CAATCCCTTGAAGAGCCCTCGGAATATGGGAACCGTCTGAAGATATGSCCTTATAT	3469
QY	3366	GGAACAACGTCAAGTCCCTGTCATGCTCAACAAGACCAATCACTATATACCAACC	3425
Db	3470	GGAATATGTTATGACCTCTTATATGTGAACATACAGACTTCTTTTAAACCGGATCC	3529
QY	3426	GGTGTGAGGCTTTGGTCTCTCAAGAAATCTTGGAGTCAACACCTGGGACGCCCAT	3485
Db	3530	TGTGTTGAACCGCTGACGCAATCTGGAACCTTGAAGCTTGAAGCCTCTCTCTAT	3589
QY	3486	CTTAAGAGGCAAGAACTGATCCGCACTGTGGCTGGGAGCAAGTGAAGCTGAACAC	3545
Db	3590	AATCAAGGGTCGGAATCTCATTCAGACAGGCCCAAGG--AATTCAGGCTTAATTAATAC	3646
QY	3546	TGTGCTGTGTTGGGAGAAACCGTGAACCGTGAACCGTGTCAAGTCAACTCTCTGGA	3605
Db	3647	AGGTGTTGATGTGACATCTCGTGTGCTTACTGTCTCCGAACCAAGTCTCTGTGGA	3706
QY	3606	GTCGCCCAACTTCATCGCAGGCAAAAGTATGACCGGTGTGTGGTGAATGAAGTCTC	3665
Db	3707	GTCTCGGAACCTTATCTGACACAGATAAAGTTACTTAAGCCGAGGGTTGAGTACTC	3766
QY	3666	CCGCGGGAATGATGATATGACCCTGGACAGGCCGCTAGCTGACGCGCCATGTCAAGAT	3725
Db	3767	ACCGGACACACTGCAATTTACTGGACAGCTGTCACTTGCCTGTCCATTAATGGAT	3826
QY	3726	CGAGTGTGCGGCGCTCTTCATCATTTTCAATGTGACCGGTCTCAATTCCTAATAAGC	3785

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: February 19, 2004, 13:00:02 ; Search time 1338 Seconds

(without alignments)
18069.130 Million cell updates/sec

Title: US-09-964-956-12

Perfect score: 5691
Sequence: 1 atgaaagccatgccttggaa.....gcttagacagctgaataaa 5691

Scoring table: IDENTITY NUC
Gapop 10.0, Gapext 1.0

Searched: 3373863 seqs, 2124039041 residues

Total number of hits satisfying chosen parameters: 6747726

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%

Listing first 45 summaries

Database : N Geneseq_290a04:*
1: geneseqn1980s:*
2: geneseqn1990s:*
3: geneseqn2000s:*
4: geneseqn2001s:*
5: geneseqn2002s:*
6: geneseqn2003s:*
7: geneseqn2003cs:*
8: geneseqn2003cs:*
9: geneseqn2003cs:*
10: geneseqn2004s:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	5691	100.0	5691	6	AB663437 Human CDN
2	5680.6	99.8	6367	6	AB600062 CADHP-9 C
3	2941.2	51.7	7136	7	ABX43468 Human mdd
4	2939.6	51.7	6387	6	ABNS9899 Novel hum
5	2778.4	48.8	6730	6	AB199504 Mouse lsc
6	2533.8	44.5	5895	6	ABT06294 Human NOV
7	2455.8	43.2	6826	8	ADA09914 Human rec
8	974	17.1	3914	6	ABV76746 Human ple
9	793.8	13.9	2597	8	ABX70006 CDNA epco
10	793.8	13.9	2597	8	ADA01359 Human PRO
11	793.8	13.9	2597	8	ADA43788 Human CDN
12	793.8	13.9	2597	8	ADA43556 Human CDN
13	793.8	13.9	2597	8	ADA01231 Human PRO
14	793.8	13.9	2597	8	ADA01115 Human CDN
15	793.8	13.9	2597	8	ADA43672 Human CDN
16	793.8	13.9	2597	8	ADA06934 Human CDN
17	793.8	13.9	2597	8	ADA08422 Novel hum
18	793.8	13.9	2597	8	ADBB9715 Human PRO
19	793.8	13.9	2597	8	ADBB6998 Human PRO
20	793.8	13.9	2597	8	ADBB6153 Human CDN
21	793.8	13.9	2597	9	ADBB9831 Human PRO
22	793.8	13.9	2597	9	ADBB9486 Novel hum
23	793.8	13.9	2597	9	ADBB6037 Human CDN

24	793.8	13.9	2597	9	ADC23435 Human CDN
25	793.8	13.9	2597	9	ADC26128 Human PRO
26	793.8	13.9	2597	9	ADBB04955 Human PRO
27	793.8	13.9	2597	9	ADBB1261 Human PRO
28	793.8	13.9	2597	9	ADBB8192 Human PRO
29	793.8	13.9	2597	9	ADBB95487 Human CDN
30	793.8	13.9	2597	9	ADBB06417 Human PRO
31	793.8	13.9	2597	9	ADBB3192 Human PRO
32	793.8	13.9	2597	9	ADBB8308 Human PRO
33	793.8	13.9	2597	9	ADBB90889 Human CDN
34	793.8	13.9	2597	10	ADBB51742 Human CDN
35	793.8	13.9	2597	10	ADBB1858 Human CDN
36	793.8	13.9	2597	10	ADBB37716 Human CDN
37	793.8	13.9	2597	10	ADBB37600 Human CDN
38	793.8	13.9	2597	10	ADBB95371 Human CDN
39	793.8	13.9	2597	10	ADBB8071 Human PRO
40	793.8	13.9	2597	10	ADBB76160 Human PRO
41	793.8	13.9	2597	10	ADBB39483 Human PRO
42	793.8	13.9	2597	10	ADBB4287 Human PRO
43	793.8	13.9	2597	10	ADBB39884 Human PRO
44	793.8	13.9	2597	10	ADBB19749 Human PRO
45	793.8	13.9	2597	10	ADBB7327 Human CDN

ALIGNMENTS

RESULT 1	AB663437	AB663437 standard; CDNA, 5691 BP.
XX	XX	XX
AC	XX	AB663437;
XX	XX	XX
DT	15-NOV-2002	(first entry)
XX	XX	XX
XX	XX	Human CDNA, homologous to plexin, designated NOV4.
XX	XX	XX
XX	XX	Human; gene, ss; NOV; NOVX; NOVX-associated disorder; cardiomyopathy;
XX	XX	atherosclerosis; diabetes; cell signaling; metabolic pathway;
XX	XX	cellular receptor; downstream effector; cancer; gene therapy;
XX	XX	hypertension; congenital heart defect; aortic stenosis; obesity;
XX	XX	infectious disease; anorexia; Alzheimer's disease; Parkinson's disease;
XX	XX	neurodegenerative disorder; haemophilia; dyslipidemia; vaccine;
XX	XX	haematopoietic disease; scleroderma; fertility; immunogen;
XX	XX	idiopathic thrombocytopenic purpura; graft versus host disease;
XX	XX	Crohn's disease; erythematous; scleritis; cirrhosis; autoimmune disease;
XX	XX	systemic lupus erythematosus; asthma; arthritis; psoriasis; allergy;
XX	XX	stroke; anxiety; Lesch-Nyhan syndrome; schizophrenia; cerebellar ataxia;
XX	XX	pain; alcoholism; transgenic.
OS	Homo sapiens.	
XX	XX	XX
XX	XX	WO200226826-A2.
XX	XX	XX
XX	04-APR-2002.	
XX	XX	XX
XX	XX	27-SEP-2001; 2001WO-US042336.
XX	XX	XX
XX	XX	27-SEP-2000; 2000US-0235631P.
XX	XX	27-SEP-2000; 2000US-0235631P.
XX	XX	27-SEP-2000; 2000US-0235808P.
XX	XX	27-SEP-2000; 2000US-0236064P.
XX	XX	27-SEP-2000; 2000US-0236065P.
XX	XX	27-SEP-2000; 2000US-0236066P.
XX	XX	28-SEP-2000; 2000US-0236135P.
XX	XX	03-OCT-2000; 2000US-0237434P.
XX	XX	03-OCT-2000; 2000US-0238321P.
XX	XX	05-OCT-2000; 2000US-0238321P.
XX	XX	06-OCT-2000; 2000US-0238399P.
XX	XX	06-OCT-2000; 2000US-0238399P.
XX	XX	16-MAR-2001; 2001US-0276667P.
XX	XX	31-MAY-2001; 2001US-0294823P.
XX	XX	12-JUL-2001; 2001US-0304868P.
XX	XX	26-SEP-2001; 2001US-00235631.

(CURA-) CURAGEN CORP.
 Gerlach VL, MacDougall JR, Smithson G, Millet I, Stone D,
 Gunther E, Ellemann K, Grose WM, Alsbrook JF, Lepley DM,
 Burgess CE, Padigaru W, Kekuda R, Spytek KA, Leach MD, Shimkets RA,
 WPI, 2002-499860/53.
 P-PSDB: AB878364.
 Novel isolated NOVX polypeptides and polynucleotides homologous to
 attractin, plexin, plexin-like family of proteins, useful for treating
 atherosclerosis, diabetes, cancer, Alzheimer's disease, hemophilia and
 stroke.
 Claim 8; Page 58-59; 308pp; English.

The invention discloses the isolated human polypeptides, and
 polynucleotides encoding them, that have been designated NOVX. The
 polypeptides, polynucleotides and antibodies are useful in treating or
 preventing a NOVX-associated disorder which is cardiomyopathy,
 atherosclerosis and diabetes in a human, where the disorder is related to
 cell signal processing and metabolic pathway modulation. They can also be
 used in determining the presence of, or predisposition to, a disease
 associated with altered levels of the polypeptides and polynucleotides of
 any one of the 13 sequences (NOV1-NOV13), for raising antibodies, for
 identifying an agent that binds to, or that modulates the expression or
 activity of the polypeptide, for identifying an agent which is cellular
 receptor or downstream effector, for treating or preventing a NOVX-
 associated disorder and as a pharmaceutical composition comprising the
 polypeptide, polynucleotide or the antibody. The polypeptides and
 polynucleotides are useful in diagnostic applications (e.g. as a marker
 for cancerous cells or tissue types) where their amounts are assessed, or
 for the manufacture of a medicament (e.g. gene therapy) for treating or
 preventing disorders or syndromes such as hypertension, congenital heart
 defects, aortic stenosis, obesity, infectious diseases, anorexia, cancer,
 Alzheimer's disease, Parkinson's disease, neurodegenerative disorders,
 hemophilia, dyslipidemia, hematopoietic diseases, scleroderma,
 fertility, idiopathic thrombocytopenic purpura, graft versus host
 disease, Crohn's disease, multiple sclerosis, cirrhosis, autoimmune
 diseases, systemic lupus erythematosus, asthma, arthritis, psoriasis,
 allergy, stroke, anxiety, Lesch-Nyhan syndrome, schizophrania, cerebellar
 ataxia, pain and alcoholism. They may also be used as immunogens to
 produce antibodies specific for the invention, and as vaccines.
 Transgenic cells containing a NOVX expressing construct are useful to
 produce non-human transgenic animals for studying the function and/or
 activity of the NOVX proteins and for identifying and/or evaluating
 modulators of NOVX protein activity. Transgenic cells containing a NOVX
 expressing construct are useful to produce non-human transgenic animals
 for studying the function and/or activity of the NOVX proteins and for
 identifying and/or evaluating modulators of NOVX protein activity. The
 sequences presented in AB863431-AB863444 are the human NOV1-NOV13 cDNAs

Sequence 5691 BP; 1318 A; 1658 C; 1574 G; 1141 T; 0 U; 0 Other;

Query Match 100.0%; Score 5691; DB 6; Length 5691;
 Best Local Similarity 100.0%; Pred. No. 0;
 Matches 5691; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

1 ATGAAGCAGATCCCTGGAACCTGACCTGCTCTCCCACTCTCATGATGAGGATG 60
 1 ATGAAGCAGATCCCTGGAACCTGACCTGCTCTCCCACTCTCATGATGAGGATG 60
 61 GGCCTCTGCACTTTGCTGACCCGAGAGCAGCCCGCTGCTCCGAAGAGCGGTATT 120
 61 GGCCTCTGCACTTTGCTGACCCGAGAGCAGCCCGCTGCTCCGAAGAGCGGTATT 120
 121 GTCAACATTCGAGAGAGCGCCGAGAGGTTTCAATCACTGATGATGAGAGACA 180
 121 GTCAACATTCGAGAGAGCGCCGAGAGGTTTCAATCACTGATGATGAGAGACA 180
 181 GGAACATTACTTGGGGGCGCTCAATGGATTAAAGCTCTCCAGAGCGAGAGGTC 240

181 GGAACATTACTTGGGGGCGCTCAATGGATTAAAGCTCTCCAGAGCGAGAGGTC 240
 241 TTGGTACGATGAGACAGGAGCCGAGAGAGCAACCCCAAGTGTATCCACCCGATC 300
 241 TTGGTACGATGAGACAGGAGCCGAGAGAGCAACCCCAAGTGTATCCACCCGATC 300
 301 GTCCAGACTGCAATGAGCCCTGACACCAACCAATGTATCAACAAATGCTCTATA 360
 301 GTCCAGACTGCAATGAGCCCTGACACCAACCAATGTATCAACAAATGCTCTATA 360
 361 GACTACAGAGAGAAAGAGCTGATGCTGAGAGCTGTACCAAGAGATGCAAGCTG 420
 361 GACTACAGAGAGAAAGAGCTGATGCTGAGAGCTGTACCAAGAGATGCAAGCTG 420
 421 CTGAGGCTGAGAGACCTCTTCAAGCTGGGGAGCCCTTATCAATGAAGAGACATCTG 480
 421 CTGAGGCTGAGAGACCTCTTCAAGCTGGGGAGCCCTTATCAATGAAGAGACATCTG 480
 481 TCAGGTGTCAACAGAGAGCGCTCACTTGTGAGATGATGCTCTTCAAGAGAGCTGAT 540
 481 TCAGGTGTCAACAGAGAGCGCTCACTTGTGAGATGATGCTCTTCAAGAGAGCTGAT 540
 541 GACAGCTGTTTCAATGTCACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 600
 541 GACAGCTGTTTCAATGTCACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 600
 601 AGCCGAGAACTGACCAAGAACTCTGAGCGAGATGATGCTGAGAGAGCTCTTCAATGAT 660
 601 AGCCGAGAACTGACCAAGAACTCTGAGCGAGATGATGCTGAGAGAGCTCTTCAATGAT 660
 661 GAGTCTGCTGCTGAGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 720
 661 GAGTCTGCTGCTGAGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 720
 721 GATATCTACTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 780
 721 GATATCTACTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 780
 781 CCGTAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 840
 781 CCGTAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 840
 841 GTGAGCTTTGCAAGAGAGAGACAGAGCTTCAATCTGATGATGATGATGATGATGAT 900
 841 GTGAGCTTTGCAAGAGAGAGACAGAGCTTCAATCTGATGATGATGATGATGATGAT 900
 901 GAGGCGATGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 960
 901 GAGGCGATGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 960
 961 GTGCTTGGAGAGAGCCCTTGAAGTCAATGATGATGATGATGATGATGATGATGATGAT 1020
 961 GTGCTTGGAGAGAGCCCTTGAAGTCAATGATGATGATGATGATGATGATGATGATGAT 1020
 1021 AAGGCGCAGAGAGCGAG 1080
 1021 AAGGCGCAGAGAGCGAG 1080
 1081 AAGCAGATTAATGACCGCATTAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1140
 1081 AAGCAGATTAATGACCGCATTAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1140
 1141 CTGACCTGAGCTGAGCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1200
 1141 CTGACCTGAGCTGAGCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1200
 1201 GACATTAATCTTGTGAGCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1260
 1201 GACATTAATCTTGTGAGCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1260
 1261 GGAATTCCTGCTTCAAG 1320
 1261 GGAATTCCTGCTTCAAG 1320

OY	132	AAGAACA	CTCTGGCCCTTGTGGGAC	CAAAAGTGGCAAGCTGAAGAA	GATCCGGG	1380	
Db	1321	AAGAACA	CTCTGTGGCTTTGTGGAC	CCAAAGTGGCAAGCTGAAGAA	GATCCGGG	1380	
OY	1381	GATGAC	CCCAAGGGCAAGCCCTCCAGAT	ATAGACGATGCAAGGTGTGA	ACCCCGGCCA	1440	
Db	1381	GATGAC	CCCAAGGGCAAGCCCTCCAGAT	ATAGACGATGCAAGGTGTGA	ACCCCGGCCA	1440	
OY	1441	GTCTCCGGGATATG	GCCTTCTCCAAAGAC	CCAGAACCTTACATCAT	ATGTCAGAGAG	1500	
Db	1441	GTCTCCGGGATATG	GCCTTCTCCAAAGAC	CCAGAACCTTACATCAT	ATGTCAGAGAG	1500	
OY	1501	CAGCTCA	CCAGAGTCCCTGTGAG	ATCCTGTGATCA	AGCTGCGGAGTGCCTT	1560	
Db	1501	CAGCTCA	CCAGAGTCCCTGTGAG	ATCCTGTGATCA	AGCTGCGGAGTGCCTT	1560	
OY	1561	GGCTCAG	GGCAACCCCACTGTGGTGTGTGTG	GTGTCGAACAACGTCGA	CCCGGAAGAG	1620	
Db	1561	GGCTCAG	GGCAACCCCACTGTGGTGTGTGTG	GTGTCGAACAACGTCGA	CCCGGAAGAG	1620	
OY	1621	CGGTGTG	AGCGTCCAAAGAGCCCGCAG	TTTCCCTCGAGAT	GAAGCAGTGTGTCGG	1680	
Db	1621	CGGTGTG	AGCGTCCAAAGAGCCCGCAG	TTTCCCTCGAGAT	GAAGCAGTGTGTCGG	1680	
OY	1681	CTGACGGT	CCATCCCAACAATATC	TCGCGATCTCAGT	ACAACGATGCGCGCTCGAG	1740	
Db	1681	CTGACGGT	CCATCCCAACAATATC	TCGCGATCTCAGT	ACAACGATGCGCGCTCGAG	1740	
OY	1741	ACGTACA	TGTCGGAGCTGTCA	GCTTGAGTCA	ACTTTGAGACCTGTCAAG	1800	
Db	1741	ACGTACA	TGTCGGAGCTGTCA	GCTTGAGTCA	ACTTTGAGACCTGTCAAG	1800	
OY	1801	ATGAGT	TGGGTGTGCTGAGCA	ATCAGATCCAGGCT	ATCCCTCGAGCAAGAGAGG	1860	
Db	1801	ATGAGT	TGGGTGTGCTGAGCA	ATCAGATCCAGGCT	ATCCCTCGAGCAAGAGAGG	1860	
OY	1861	CCCCGAT	CAACAAGATGGGAG	CCACCATGTGTCA	GAGTTCAGCTCAAAATCAAG	1920	
Db	1861	CCCCGAT	CAACAAGATGGGAG	CCACCATGTGTCA	GAGTTCAGCTCAAAATCAAG	1920	
OY	1921	GAGACGG	CAATGACCTTGGCCAG	CAACGACTTGTCTTAC	CAATTGCAGACGCTCAACAT	1980	
Db	1921	GAGACGG	CAATGACCTTGGCCAG	CAACGACTTGTCTTAC	CAATTGCAGACGCTCAACAT	1980	
OY	1981	TGCTGCTGT	CTGTGGTGAAGT	CAATACCGCTG	CACTGTGTATATCCGACATGTC	2040	
Db	1981	TGCTGCTGT	CTGTGGTGAAGT	CAATACCGCTG	CACTGTGTATATCCGACATGTC	2040	
OY	2041	TGACCCAT	TGACCCCAAGACCTG	CTCTCCAGAAAGCCAG	TGAAGCTGCGCCGAGAC	2100	
Db	2041	TGACCCAT	TGACCCCAAGACCTG	CTCTCCAGAAAGCCAG	TGAAGCTGCGCCGAGAC	2100	
OY	2101	TGCCCCAG	TGCTGCGATGAG	CAAGATCCTGTGTCC	CGTGAAGTATCAAGCTATC	2160	
Db	2101	TGCCCCAG	TGCTGCGATGAG	CAAGATCCTGTGTCC	CGTGAAGTATCAAGCTATC	2160	
OY	2161	ACGCTGA	AGGCCAAGAACTCCCGAG	CCCAAGTCTGGAG	CAGGTGGCTTACAAATGATC	2220	
Db	2161	ACGCTGA	AGGCCAAGAACTCCCGAG	CCCAAGTCTGGAG	CAGGTGGCTTACAAATGATC	2220	
OY	2221	CTCAACAT	TGAGGCAAGCAG	CAGCAGTGTCCGCTGTG	CTTCAACGCTCAAGCTT	2280	
Db	2221	CTCAACAT	TGAGGCAAGCAG	CAGCAGTGTCCGCTGTG	CTTCAACGCTCAAGCTT	2280	
OY	2281	CAGTGC	CAAGAACCTCTTAT	TCATCAAGGA	TGAGATCAACAACCTG	CCCGTGAAG	2340
Db	2281	CAGTGC	CAAGAACCTCTTAT	TCATCAAGGA	TGAGATCAACAACCTG	CCCGTGAAG	2340
OY	2341	TTGACAG	TGCTGTGGAATGG	GCATTTCAACTTGA	CAACCCAGCTCAAGATTAAGTTCA	2400	
Db	2341	TTGACAG	TGCTGTGGAATGG	GCATTTCAACTTGA	CAACCCAGCTCAAGATTAAGTTCA	2400	

QY	2401	TTCTAACAAGTGTGGAGGCCATCGTGGTGAAGCTGACGGGCTGTGCTCTCAAGAGCTGACCCAGAC	2458
Db	2401	CTTTCACAACTGTGAAGCCATCTCGTGAAGAGCTCGGGCTGTGCTCTCAAGAGCTGACCCAGAC	2466
QY	2461	TTTCGATGTGGCTGTGTGCAGAGGCCCAAGCCATGTGCACCTTGTCCGCACACTCTGCCCTGTGC	2520
Db	2461	TTTCGATGTGTGCTGTGTGTGCAGAGGCCCAAGCCATGTGTGCACCTTGTCCGCACACTCTGCCCTGTGC	2520
QY	2521	CAGAGAGCCAGTGTGAGTGTGAGAGCTGTCTGTGTGTCTCAAAAGCAAGTGTCACAAAACCCCGCATC	2580
Db	2521	CAGAGAGCCAGTGTGAGTGTGAGAGCTGTCTGTGTGTCTCAAAAGCAAGTGTCACAAAACCCCGCATC	2580
QY	2581	ACAGAGATATATCCCGGTGAACAGGCCCCCGGGAAGGGGACCAAGGTCACTATCCGAGGG	2640
Db	2581	ACAGAGATATATCCCGGTGAACAGGCCCCCGGGAAGGGGACCAAGGTCACTATCCGAGGG	2640
QY	2641	GAGAACCTGGGACCTGTGAATTTCCGACATGAGCTCCCATGTCAAGGTTGTGCGGTGGAG	2700
Db	2641	GAGAACCTGGGACCTGTGAATTTCCGACATGAGCTCCCATGTCAAGGTTGTGCGGTGGAG	2700
QY	2701	TGCAGCCCTTTAGTGTGATGTGTTACATCCCTGTACAGAACATCGTGTGTGAGATGGGGAG	2760
Db	2701	TGCAGCCCTTTAGTGTGATGTGTTACATCCCTGTACAGAACATCGTGTGTGAGATGGGGAG	2760
QY	2761	GCCAGAGCCAGCCAGACATGACAGAGCTCTGTGGAGATGTGCGGTGTGTGTGTGCGGCTGTAA	2820
Db	2761	GCCAGAGCCAGCCAGACATGACAGAGCTCTGTGGAGATGTGCGGTGTGTGTGTGCGGCTGTAA	2820
QY	2821	TTTCATGGCCCGGTCTCTCAAGCTCTATTTACTTTCATGACATGACATCTCTCTCAAGTCTGAG	2880
Db	2821	TTTCATGGCCCGGTCTCTCAAGCTCTATTTACTTTCATGACATGACATCTCTCTCAAGTCTGAG	2880
QY	2881	CCCAAGCCGGGGGCCCATGTCCGAGAGGAGCCCAATGTACATCAAGGACCCAACTTGAAT	2940
Db	2881	CCCAAGCCGGGGGCCCATGTCCGAGAGGAGCCCAATGTACATCAAGGACCCAACTTGAAT	2940
QY	2941	GCCGGAGCAACGTGTGTGTGTATGTTTGGAAAGCAGCCCTGTCTCTTCCACAGGCCCATCT	3000
Db	2941	GCCGGAGCAACGTGTGTGTGTATGTTTGGAAAGCAGCCCTGTCTCTTCCACAGGCCCATCT	3000
QY	3001	CCATCCCAATTTGTCTGCAACACACACATCCCTCAGATGAGGTGTGAGATGAAAGGTGTG	3060
Db	3001	CCATCCCAATTTGTCTGCAACACACACATCCCTCAGATGAGGTGTGAGATGAAAGGTGTG	3060
QY	3061	GTGCAGTGTGACAGAGGCCAAGATCCACCAAGACCTGGTCTTTCAAGTATGTGTGAAGAAGCC	3120
Db	3061	GTGCAGTGTGACAGAGGCCAAGATCCACCAAGACCTGGTCTTTCAAGTATGTGTGAAGAAGCC	3120
QY	3121	ACCATTCGTGGAGTTTGAAGCAGATGTGAGAGCTGTGCTGAGGAACCAACCCATGTCCGCTA	3180
Db	3121	ACCATTCGTGGAGTTTGAAGCAGATGTGAGAGCTGTGCTGAGGAACCAACCCATGTCCGCTA	3180
QY	3181	TGGGGAGACCCACTGTGAACCTCATACAGAACCCCAATCCGTGTGCAAGCATGTGAGGGAG	3240
Db	3181	TGGGGAGACCCACTGTGAACCTCATACAGAACCCCAATCCGTGTGCAAGCATGTGAGGGAG	3240
QY	3241	GAGACATCATATATCTGTGAGGTTCTGAACGCTTCTAGATGTGACCTGTGACGCGCCGCG	3300
Db	3241	GAGACATCATATATCTGTGAGGTTCTGAACGCTTCTAGATGTGACCTGTGTGACGCGCCGCG	3300
QY	3301	CTCGCTCTGGGTCTGTACACACAGTCAAGCTGTGACCGAAGAGGCCCGAGAGATTGTGGCTTC	3360
Db	3301	CTCGCTCTGGGTCTGTGTACACACAGTCAAGCTGTGACCGAAGAGAGGCCCGAGAGATTGTGGCTTC	3360
QY	3361	ATCTGTGACAACTGTCAAGTCCCTGTCTATCTCACTCAACAAAGCACTTCACTACTATATCC	3420
Db	3361	ATCTGTGACAACTGTCAAGTCCCTGTCTATCTCACTCAACAAAGCACTTCACTACTATATCC	3420
QY	3421	AACCCGCTGTGTAGAGCTTTGTGTCCTCAAGAAATCTGTGAGCTCAAGCCTGTGTGAGAGCC	3480
Db	3421	AACCCGCTGTGTAGAGCTTTGTGTCCTCAAGAAATCTGTGAGCTCAAGCCTGTGTGAGAGCC	3480
QY	3481	ATCATCTCTAAAGGCAAGACTGTATCCCGCTGTGTGCTGTGGGGCAACGTGAACCTGAC	3540

```
Db 3481 ATCATCTTAAGGCGAAGAACTGATCCCGCTGTGGCTGGGGGCAACTGAAAGCTGAAC 3540
Qy 3541 TACACTGTGTGTGGTGGAGAAAGCGGTGCAACGTGACCGTGTCAATGTCCAGTCTGTC 3600
Db 3541 TACACTGTGTGTGGTGGAGAAAGCGGTGCAACGTGACCGTGTCAATGTCCAGTCTGTC 3600
Qy 3601 TGGAGTGTCCCAACTCATCGAGGCAAGATGATGAGCCGTGTGGTGGATGAG 3660
Db 3601 TGGAGTGTCCCAACTCATCGAGGCAAGATGATGAGCCGTGTGGTGGATGAG 3660
Qy 3661 TACTTCCCGGGGATGTGTATTTGCCCCGAGACGCCCTGACGCTCCCGCATGTC 3720
Db 3661 TACTTCCCGGGGATGTGTATTTGCCCCGAGACGCCCTGACGCTCCCGCATGTC 3720
Qy 3721 AGCATCGAGTGTGGCGGCTCTCATATTTCATCGTGGCCGTGTCAATGGCTAT 3780
Db 3721 AGCATCGAGTGTGGCGGCTCTCATATTTCATCGTGGCCGTGTCAATGGCTAT 3780
Qy 3781 AAAGCGAAGTCCCGGAAAGTGAACCTGACCGTGAAGGGGTGCAATGACATGACAC 3840
Db 3781 AAAGCGAAGTCCCGGAAAGTGAACCTGACCGTGAAGGGGTGCAATGACATGACAC 3840
Qy 3841 CTGGAGTCCCGTGTGGCCCTGAGTGAAGAAAGCTTTGCGAGCTGACAGACGACATC 3900
Db 3841 CTGGAGTCCCGTGTGGCCCTGAGTGAAGAAAGCTTTGCGAGCTGACAGACATC 3900
Qy 3901 CATAGACTGACCAATGACCTGATGAGACCGGGATTCCTGCTGACATAGAACTTAC 3960
Db 3901 CATAGACTGACCAATGACCTGATGAGACCGGGATTCCTGCTGACATAGAACTTAC 3960
Qy 3961 ACCATGCGGGTGTGTGCTCCAGAAATTGAAGACCAACCTGTCTCCGGAGCTTGAGTC 4020
Db 3961 ACCATGCGGGTGTGTGCTCCAGAAATTGAAGACCAACCTGTCTCCGGAGCTTGAGTC 4020
Qy 4021 CCGGGCTACCGGAGAGCGGTGTGAGAAAGGCTGAAAGCTCTTCCGACCACTCATCAC 4080
Db 4021 CCGGGCTACCGGAGAGCGGTGTGAGAAAGGCTGAAAGCTCTTCCGACCACTCATCAC 4080
Qy 4081 AACAGGTGTCTGTGTCTCTTCATCCGACCGCTTGAAGTCCAGAGCTGATCTTCATG 4140
Db 4081 AACAGGTGTCTGTGTCTCTTCATCCGACCGCTTGAAGTCCAGAGCTGATCTTCATG 4140
Qy 4141 CGGACCGTGGCAAGTGTGCTCATCATGACCGTGTGAGACCAAGCTGAGATC 4200
Db 4141 CGGACCGTGGCAAGTGTGCTCATCATGACCGTGTGAGACCAAGCTGAGATC 4200
Qy 4201 GCCACTGATGTGTGAAGCAAGCTGTGCGGACCTCATTTGACAAAGAACTGAGACAG 4260
Db 4201 GCCACTGATGTGTGAAGCAAGCTGTGCGGACCTCATTTGACAAAGAACTGAGACAG 4260
Qy 4261 AACCAACCTGAGCTGTGTGAGAGCAAGTGAAGATGAGTGTGAGAAAGTGTGAGAT 4320
Db 4261 AACCAACCTGAGCTGTGTGAGAGCAAGTGAAGATGAGTGTGAGAAAGTGTGAGAT 4320
Qy 4321 TGGTTACTTCTCTCTCAATGATTCCTCAAGAGTGTCTGGGAGCCCTCTCTTCAC 4380
Db 4321 TGGTTACTTCTCTCTCAATGATTCCTCAAGAGTGTCTGGGAGCCCTCTCTTCAC 4380
Qy 4381 CTGTTCTGTGCAATCAAGACAGATGAGAAAGGCGCCATTTGACCGCATCAGGAGAG 4440
Db 4381 CTGTTCTGTGCAATCAAGACAGATGAGAAAGGCGCCATTTGACCGCATCAGGAGAG 4440
Qy 4441 GCCCGCTACTCTTTGAGGAGAGCAAGCTATCCGACAGATTTGATCAAAACCTTG 4500
Db 4441 GCCCGCTACTCTTTGAGGAGAGCAAGCTATCCGACAGATTTGATCAAAACCTTG 4500
Qy 4501 GTCCGAGCTGTGTCAAGCCAGACATGTCCAAAGCCGAGGTCTCCAGTAAAGATCTC 4560
Db 4501 GTCCGAGCTGTGTCAAGCCAGACATGTCCAAAGCCGAGGTCTCCAGTAAAGATCTC 4560
Qy 4561 AACTGTGACCAATCACTGAGTGAAGAGATTTTGGATGCACTTTCAAGAAATGTG 4620
Db 4561 AACTGTGACCAATCACTGAGTGAAGAGATTTTGGATGCACTTTCAAGAAATGTG 4620

Db 4621 AACTGTGACCAATCACTGAGTGAAGAGATTTTGGATGCACTTTCAAGAAATGTG 4680
Qy 4621 CTTTGTCTCCACCGGCGCAAAAGTGTGAGATTTGAGTGTGGGCAAGAAAGTGG 4680
Db 4621 CTTTGTCTCCACCGGCGCAAAAGTGTGAGATTTGAGTGTGGGCAAGAAAGTGG 4680
Qy 4681 GCAAGATGATCTTGGAGATGAAGATCAACCAAGATTTGAGATTTGAGAGCA 4740
Db 4681 GCAAGATGATCTTGGAGATGAAGATCAACCAAGATTTGAGATTTGAGAGCA 4740
Qy 4741 CTGAACACTGGCCCACTACAGATGCAAGTGTGCGGTGGGCAATTTGATGTCAG 4800
Db 4741 CTGAACACTGGCCCACTACAGATGCAAGTGTGCGGTGGGCAATTTGATGTCAG 4800
Qy 4801 CAGGTGACGCTTATACGAGTGAACCACTCACCGTCTCCAGAGCTTCAGAAATTA 4860
Db 4801 CAGGTGACGCTTATACGAGTGAACCACTCACCGTCTCCAGAGCTTCAGAAATTA 4860
Qy 4861 TATGAATAATGATCCGATACAGGCGAGCCCGGACCGCTCCGTCAGAGACCTATG 4920
Db 4861 TATGAATAATGATCCGATACAGGCGAGCCCGGACCGCTCCGTCAGAGACCTATG 4920
Qy 4921 ATCACTCTGACCTGAGAGTGAAGTCAAGATGTGGCACTTGTGAAGAACCAAGAC 4980
Db 4921 ATCACTCTGACCTGAGAGTGAAGTCAAGATGTGGCACTTGTGAAGAACCAAGAC 4980
Qy 4981 GGAACCAAGAGAGGAGGAGCCGGGGAGCAAGTGTGTGAATCTACTGACCCGA 5040
Db 4981 GGAACCAAGAGAGGAGGAGCCGGGGAGCAAGTGTGTGAATCTACTGACCCGA 5040
Qy 5041 CTCCTGGCACTAAGGCACTGACAAAGTGTGTGATGACCTGTTGAGACATCTTC 5100
Db 5041 CTCCTGGCACTAAGGCACTGACAAAGTGTGTGATGACCTGTTGAGACATCTTC 5100
Qy 5101 AGAAGGCAACCGTGTGTGCTGCTGCGGCGGCGGCACTGATGATTTGATGCTGTG 5160
Db 5101 AGAAGGCAACCGTGTGTGCTGCTGCGGCGGCGGCACTGATGATTTGATGCTGTG 5160
Qy 5161 GATGAGCAGCTGATTAACATGATGATGACCGGCAAGTGTGATGACCTGTTGAGAGC 5220
Db 5161 GATGAGCAGCTGATTAACATGATGATGACCGGCAAGTGTGATGACCTGTTGAGAGC 5220
Qy 5221 AATTGCTGCGCTGAGTGTGAGTGTGAGTCAATGATGAAGAACCGGATTTGTGTGAC 5280
Db 5221 AATTGCTGCGCTGAGTGTGAGTGTGAGTCAATGATGAAGAACCGGATTTGTGTGAC 5280
Qy 5281 ATCCATTAAGAAAGATCAAGACCGCTGTCTGTGTGTGTGATGACCTTCAATGAC 5340
Db 5281 ATCCATTAAGAAAGATCAAGACCGCTGTCTGTGTGTGTGATGACCTTCAATGAC 5340
Qy 5341 TCTTGTCCACGTCAGAGACCGGCTGGGCAAGTGTGAGTGTGATGACCTTCAATGAC 5400
Db 5341 TCTTGTCCACGTCAGAGACCGGCTGGGCAAGTGTGAGTGTGATGACCTTCAATGAC 5400
Qy 5401 GCCAAGACATCCCGGCTCAAGAAAGTGTGTGAGAGTATTAATGACATGAGAGAG 5460
Db 5401 GCCAAGACATCCCGGCTCAAGAAAGTGTGTGAGAGTATTAATGACATGAGAGAG 5460
Qy 5461 ATCCAGGACATGAGGACCAAGATGAGTGAAGTGTGAGAGTATTAATGAGAGAGAG 5520
Db 5461 ATCCAGGACATGAGGACCAAGATGAGTGAAGTGTGAGAGTATTAATGAGAGAGAG 5520
Qy 5521 ATGAATGATTTCAACCAATGATGATGATGATGATGATGATGATGATGATGATGAT 5580
Db 5521 ATGAATGATTTCAACCAATGATGATGATGATGATGATGATGATGATGATGATGAT 5580
Qy 5581 AGGAGAGATCTTGTGACCTGTGACCAAGTGAACAGTGTGAGAGCAAGAACTGGCC 5640
Db 5581 AGGAGAGATCTTGTGACCTGTGACCAAGTGAACAGTGTGAGAGCAAGAACTGGCC 5640
Qy 5641 TACAACTGAACCAAGTCAATGATGATGATGATGATGATGATGATGATGATGATGAT 5691
Db 5641 TACAACTGAACCAAGTCAATGATGATGATGATGATGATGATGATGATGATGATGAT 5691
```


1646 AAGCAGTAAATGACCCGATTAAAGACCGGCTGAGTCTTTGTTACCGGAGGCAAGGGACG 1705
Qy 1141 CTGACCTGGGCTGGCTCAAGGTGAAGAGACATCCCTGACAGAGTGGCTCTTAACATT 1200
Db 1706 CTGGACCTGGGCTGGCTCAAGGTGAAGAGACATCCCTGACAGAGTGGCTCTTAACATT 1765
Qy 1201 GACGATTAATCTCTGAGCTGAGACATGATGCTCCCTGGAGATGACCAATGGTGGCT 1260
Db 1766 GACGATTAATCTCTGAGCTGAGACATGATGCTCCCTGGAGATGACCAATGGTGGCT 1825
Qy 1261 GGAATTCCTGCTTCAAGGAGACAGGAGCCGATGACGCTGATCTGATCTGATCTGATC 1320
Db 1826 GGAATTCCTGCTTCAAGGAGACAGGAGCCGATGACGCTGATCTGATCTGATCTGATC 1885
Qy 1321 AAGAACCACTCTGAGGCTTGTGGGACCAAAAGTGGCAAGCTGAAGAAAGATCCGGAGT 1380
Db 1886 AAGAACCACTCTGAGGCTTGTGGGACCAAAAGTGGCAAGCTGAAGAAAGATCCGGAGT 1945
Qy 1381 GATGAGCCAGGAGCAAGCCCTTCAGTATGAGACGCTGAGGCTGATGATCCCGGACCA 1440
Db 1946 GATGAGCCAGGAGCAAGCCCTTCAGTATGAGACGCTGAGGCTGATGATCCCGGACCA 2005
Qy 1441 GTCCCTCCGAGATGAGCTTCTCCAGAGACCAAGAGCACTTACATCTGATCAAGAGAG 1500
Db 2006 GTCCCTCCGAGATGAGCTTCTCCAGAGACCAAGAGCACTTACATCTGATCAAGAGAG 2065
Qy 1501 CAGCTACCAAGATCTCTGAGAGTCTGATGATGATGATGATGATGATGATGATGATGAT 1560
Db 2066 CAGCTACCAAGATCTCTGAGAGTCTGATGATGATGATGATGATGATGATGATGATGAT 2125
Qy 1561 GAGTCAAGGAGCAAGCCCTTCAGTATGAGACGCTGAGGCTGATGATCCCGGACCA 1620
Db 2126 GAGTCAAGGAGCAAGCCCTTCAGTATGAGACGCTGAGGCTGATGATCCCGGACCA 2185
Qy 1621 CGGTGTGAGGCTGATGAGAGCCCGGAGGTTGGCTCGAGATGAGAGTGTGTCCG 1680
Db 2186 CGGTGTGAGGCTGATGAGAGCCCGGAGGTTGGCTCGAGATGAGAGTGTGTCCG 2245
Qy 1681 CTGACGCTGATCCCAACATATCTCCGCTCTCAGATCAAGCTGCTGCTGCTGAG 1740
Db 2246 CTGACGCTGATCCCAACATATCTCCGCTCTCAGATCAAGCTGCTGCTGAG 2305
Qy 1741 ACGTACATATCTCCGAGCTGTCACTGAGCTCACTGAGCCTTTGAGAGCTGTGAG 1800
Db 2306 ACGTACATATCTCCGAGCTGTCACTGAGCTCACTGAGCCTTTGAGAGCTGTGAG 2365
Qy 1801 ATGATGAGGCTGATGAGGCAATCAGATCAGTGTACTCCCTGACAGCAGAGAGT 1860
Db 2366 ATGATGAGGCTGATGAGGCAATCAGATCAGTGTACTCCCTGACAGCAGAGAGT 2425
Qy 1861 CCCCAGATCATCAGAGAAATGGGAGACCAATGTGTGATGATGATGATGATGATGAT 1920
Db 2426 CCCCAGATCATCAGAGAAATGGGAGACCAATGTGTGATGATGATGATGATGATGAT 2485
Qy 1921 GAGACCGGAGATGAGCTTGGCAGGACCAAGCTTTGCTTCAATTTGACAGCTTCAAT 1980
Db 2486 GAGACCGGAGATGAGCTTGGCAGGACCAAGCTTTGCTTCAATTTGACAGCTTCAAT 2545
Qy 1981 TGGTACCTGCTGCTGATGAGAGTGCATACCGGCTGACAGTGTGATTAATTCGGGAT 2040
Db 2546 TGGTACCTGCTGCTGATGAGAGTGCATACCGGCTGACAGTGTGATTAATTCGGGAT 2605
Qy 2041 TGCACCATGACCCCAAGACCTGCTCTTCCAGAGAGCCGAGTGAAGCTGCCGAGAG 2100
Db 2606 TGCACCATGACCCCAAGACCTGCTCTTCCAGAGAGCCGAGTGAAGCTGCCGAGAG 2665
Qy 2101 TGGCCCGAGCTGCTGAGTGAAGTGAAGTCTGATGAGCCGAGAGGATCAAGCTTAC 2160
Db 2666 TGGCCCGAGCTGCTGAGTGAAGTGAAGTCTGATGAGCCGAGAGGATCAAGCTTAC 2725
Qy 2161 AGCTGAAGGCGCAAGACCTTCCCGAGCCCGAGCTGTGGGACGCTGATCAAGATCATC 2220

2726 ACGCTGAAGGCCAAGAACTCCCCAGGCCAGGTGTGGGCAAGCGTGTACGAATGATC 2785
Qy 2221 CTCAATATTCAGGAGCAGAGACAGAGAGTGGCCCGCTGCTGCTTCAACAGCTCCAGCTA 2280
Db 2786 CTCAATATTCAGGAGCAGAGACAGAGAGTGGCCCGCTGCTGCTTCAACAGCTCCAGCTA 2845
Qy 2281 CAGTGCAGAAACCTCTTATTTCTATGAGAGGATGAGATCAACACTGCTCCGTGGAG 2340
Db 2846 CAGTGCAGAAACCTCTTATTTCTATGAGAGGATGAGATCAACACTGCTCCGTGGAG 2905
Qy 2341 TTGACGTGTGTGGAATGGGCACTTCAACATTTGACCAACCCAGCTCAGAAATTAAGTAC 2400
Db 2906 TTGACGTGTGTGGAATGGGCACTTCAACATTTGACCAACCCAGCTCAGAAATTAAGTAC 2965
Qy 2401 CTCTAAGGTGAGACCAAGGAGAGCTGCGGAGCTGCTGCTCAAGCTGACCCAGAC 2460
Db 2966 CTCTAAGGTGAGACCAAGGAGAGCTGCGGAGCTGCTGCTCAAGCTGACCCAGAC 3025
Qy 2461 TTGGCATGTGGCTGGGCGAGGAGCCAGGCGAGTGCACCTGGCGCAGACCTGGCC 2520
Db 3026 TTGGCATGTGGCTGGGCGAGGAGCCAGGCGAGTGCACCTGGCGCAGACCTGGCC 3085
Qy 2521 CAGAGAGCCAGTGTGAGAGCTGTGTGTGTCGCAAAAGCAAGTGCACAAACCCCGATC 2580
Db 3086 CAGAGAGCCAGTGTGAGAGCTGTGTGTGTCGCAAAAGCAAGTGCACAAACCCCGATC 3145
Qy 2581 ACAGAGATATATCCCGGTGACAGGCCCCCGGAGAGGGGAGCCAGAGTCACTATCCAGAG 2640
Db 3146 ACAGAGATATATCCCGGTGACAGGCCCCCGGAGAGGGGAGCCAGAGTCACTATCCAGAG 3205
Qy 2641 GAGAACCTGGGCTGAGATTTGCGACATGCTGCTCCATGTCAAGTGTGCGGTGAG 2700
Db 3206 GAGAACCTGGGCTGAGATTTGCGACATGCTGCTCCATGTCAAGTGTGCGGTGAG 3265
Qy 2701 TGCAGGCTTTGATGAGATGATTAATCCCTGAGAGACAGTGTGTGAGTGGGAGAG 2760
Db 3266 TGCAGGCTTTGATGAGATGATTAATCCCTGAGAGACAGTGTGTGAGTGGGAGAG 3325
Qy 2761 GCCAAGCCCAAGCAGATCAGGCTTTGAGAGATCTGAGGATGTGTGTGCTGAA 2820
Db 3326 GCCAAGCCCAAGCAGATCAGGCTTTGAGAGATCTGAGGATGTGTGTGCTGAA 3385
Qy 2821 TTGATGGCCCGTGTCCAGAGCTCTATTACTTATGACAGTACTCTGATGTGAG 2880
Db 3386 TTGATGGCCCGTGTCCAGAGCTCTATTACTTATGACAGTACTCTGATGTGAG 3445
Qy 2881 CCAAGCCGAGGAGCCATGCTCGAGAGGAGCCAGATGACCATCAAGGACCAACTGAT 2940
Db 3446 CCAAGCCGAGGAGCCATGCTCGAGAGGAGCCAGATGACCATCAAGGAGCAACTGAT 3505
Qy 2941 GCCGAGAGCAGTGTGTGATGTTTGAAGAGAGCCCTGCTCTTCAAGGAGTCT 3000
Db 3506 GCCGAGAGCAGTGTGTGATGTTTGAAGAGAGCCCTGCTCTTCAAGGAGTCT 3565
Qy 3001 CCAATCTCAATTTGTCTGACACACACATCTCAATGAGAGTCTGAGATGAAGTGTG 3060
Db 3566 CCAATCTCAATTTGTCTGACACACACATCTCAATGAGAGTCTGAGATGAAGTGTG 3625
Qy 3061 GTGAGGATGAGCAGGAGCCAGATCCACAGAGCTGTGCTTCAATGATGAGAGAGCC 3120
Db 3626 GTGAGGATGAGCAGGAGCCAGATCCACAGAGCTGTGCTTCAATGATGAGAGAGCC 3685
Qy 3121 ACCATGTGTGATGAGGAGATGAGAGATTTGTGATGAGAAACACACCCATGCCGTA 3180
Db 3686 ACCATGTGTGATGAGGAGATGAGAGATTTGTGATGAGAAACACACCCATGCCGTA 3745
Qy 3181 TGGGAGAGCCAGCTGAGCTCATACAGAAACCCCGAGTCCGTCAGAGATGAGAGAG 3240
Db 3746 TGGGAGAGCCAGCTGAGCTCATACAGAAACCCCGAGTCCGTCAGAGATGAGAGAG 3805
Qy 3241 GAGCACTCAATATCTGTGAGGCTTGAAGAGCTGATGAGATCACTGTGACAGGCGCCGC 3300
Db 3806 GAGCACTCAATATCTGTGAGGCTTGAAGAGCTGATGAGATCACTGTGAGGCGCCGC 3865

DB 6026 ATGCCAGCCATCAGGACCAACAGCATGAAACGATACCTGGCTGAGCAGTCCCGGATGCAC 6085
 QY 5521 ATGAATGAGTTCACACCACTGAGTGCATCTCTGAGAGATCTCTCTTGTGGGCAATAC 5580
 DB 6086 ATGAATGAGTTCACACCACTGAGTGCATCTCTGAGAGATCTCTCTTGTGGGCAATAC 6145
 QY 5581 AGCGAGAGATCTCTGAGTGCATCTCTGAGAGATCTCTCTTGTGGGCAATAC 5640
 DB 6146 AGCGAGAGATCTCTGAGTGCATCTCTGAGAGATCTCTCTTGTGGGCAATAC 6205
 QY 5641 TACAACTGAGAGATCTCTGAGTGCATCTCTGAGAGATCTCTCTTGTGGGCAATAC 5697
 DB 6206 TACAACTGAGAGATCTCTGAGTGCATCTCTGAGAGATCTCTCTTGTGGGCAATAC 6252
 RESULT 3
 ABX34698
 ID ABX34698 standard, cDNA, 7136 BP.
 AC ABX34698;
 XX
 DT 13-FEB-2003 (first entry)
 DE Human mdct cDNA SEQ ID 259.
 XX
 MMDDT: human, disease detection and treatment molecule polypeptide;
 KW anti-inflammatory; immunosuppressive; osteoporotic; cytotoxic; anti-HIV;
 KW hemostatic; nephrotoxic; antianemic; antiproliferative; hepatotropic;
 KW gene therapy; protein leukaemia; lymphoma; melanoma; myeloma; sarcoma;
 KW cancer; adenocarcinoma; leukaemia; lymphoma; melanoma; myeloma; sarcoma;
 KW anaemia; Crohn's disease; acquired immunodeficiency syndrome; AIDS;
 KW Goodpasture's syndrome; inflammation; osteoporosis; thrombocytopenia;
 KW peritasis; hepatitis; gene; ss.
 KM
 OS Homo sapiens.
 XX
 PN WO200279449-A2.
 XX
 PD 10-OCT-2002.
 XX
 PF 27-MAR-2002; 2002MO-US009944.
 XX
 PR 28-MAR-2001; 2001US-0279619P.
 PR 29-MAR-2001; 2001US-0280067P.
 PR 29-MAR-2001; 2001US-0280068P.
 PR 16-MAY-2001; 2001US-0291280P.
 PR 17-MAY-2001; 2001US-0291829P.
 PR 17-MAY-2001; 2001US-0291849P.
 PR 19-JUN-2001; 2001US-0299428P.
 PR 20-JUN-2001; 2001US-0299776P.
 PR 20-JUN-2001; 2001US-0300001P.
 PR
 PA (INCY-) INCYTE GENOMICS INC.
 XX
 PI Daffo A, Jones AL, Tran AB, Dahl CR, Gietzen D, Chinn J,
 PI Dufour GG, Hillman JL, Yu JY, Tuason O, Yap PE, Amshy SR,
 PI Dufour SC, Dam TC, Liu TF, Nguyen DA, Kleefeld Y, Gerstin EH,
 PI Peralta CH, David MH, Lewis SA, Chen AJ, Panzer SR, Harris B,
 PI Flores V, Marwaha R, Lo A, Lan RY, Urashka ME,
 XX
 DR WPI, 2003-058431/05.
 DR P-SDB; AB011708.
 XX
 PT New purified disease detection and treatment molecule proteins and
 PT polynucleotides, useful for diagnosing, treating or preventing cancers
 PT (e.g. leukemia or sarcoma), anemia, Crohn's disease, AIDS, osteoporosis
 PT or hepatitis.
 XX
 PS Claim 1, SEQ ID NO 259; 339pp + Sequence Listing, English.
 XX
 CC This invention describes a novel disease detection and treatment molecule
 CC polypeptide (MDDT) which has anti-inflammatory, immunosuppressive,

CC osteoporotic, cytostatic, anti-HIV, haemostatic, nephrotoxic,
 CC antianemic, antiproliferative and hepatotropic activity. The polynucleotides
 CC and the polypeptides of the invention can be used for gene therapy,
 CC protein replacement therapy and are useful for treating a variety of
 CC diseases or conditions. These polypeptides or polynucleotides are
 CC particularly useful for diagnosing, treating or preventing cell
 CC proliferative disorders (e.g. cancers including adenocarcinoma,
 CC leukemia, lymphoma, melanoma, myeloma or sarcoma), anaemia, Crohn's
 CC disease, acquired immunodeficiency syndrome (AIDS), Goodpasture's
 CC syndromes, inflammation, osteoporosis, thrombocytopenia, peritasis or
 CC hepatitis. ABX3440-ABX34835 encode the MDDT polypeptides represented in
 CC AB011450-AB011645, described in the disclosure of the invention. NOTE:
 CC The sequence data for this patent did not form part of the printed
 CC specification, but was obtained in electronic format from WIPO at
 CC ftp.wipo.int/pub/published_pat_sequences
 XX
 SQ Sequence 7136 BP; 1616 A; 2099 C; 1993 G; 1428 T; 0 U; 0 Other;
 XX
 Query Match 51.7%; Score 2941.2; DB 7; Length 7136;
 Best Local Similarity 71.3%; Pred. No. 0;
 Matches 3953; Conservative 0; Mismatches 1578; Indels 15; Gaps 5;
 QY 151 TTCATACCTGCTGTGATGATGAGACAGACATTTACTTGGGGCCCTCAATCG 210
 DB 853 TTCAACCACTTGAACCTGTCACCAAGGAGCGGCGCTGTATGTGGGCGCATCAACCG 912
 QY 211 ATTACAGCTCTCCAGCAGCCTGAAAGTCTTGTGTGACGATGAGACAGGCGGACGAG 270
 DB 913 GTCTATGAGCTGACAGGCACTGACCATTCAGAGTGTGCTATTAAGACAGGGCCAAAGAG 972
 QY 271 GACACCCCAAGTATACCCAGCCCGCATGTCAGACCTGGAATGAGCCCTGACGAC 330
 DB 973 GACACAGCTCTGTATACCCGCTTCACTGTGACCTTGAAGCAAGTGTACCTC 1032
 QY 331 ACCAAGATGTCAACAAGATGCTCTCATGACTAAGAGAGACAGGCTGATTCCTGT 390
 DB 1033 ACCAAGATGTCAACAAGATGCTCTCATGACTAAGAGAGACAGGCTGATTCCTGT 1092
 QY 391 GAGAGCCTGACCAAGGATCTGCAAGCTGAGAGCTGAGAGACCTCTCAAGTGTGG 450
 DB 1093 GAGAGCCTGACCAAGGATCTGCAAGCTGAGAGCTGAGAGACCTCTCAAGTGTGG 1152
 QY 451 GAGCCTTATCATTAAGAGAGACATCTGTGAGGTGTACAGAGAGGCTGACTCTT 510
 DB 1153 GAGCCTTATCATTAAGAGAGACATCTGTGAGGTGTACAGAGAGGCTGACTCTT 1212
 QY 511 GAGTGTATGCTCTCTACAGCAACTGATGACAGCTGTTTATGACGAGAGTGTGAT 570
 DB 1213 GAGGTATGTGGCTCTGAGAGGTGAGATGAGCAAGCTTCTATGAGAGGCTGTGAT 1272
 QY 571 GGGAGCCCGATATTTTCCACATCTCCAGCCGGAATGACCAAGACTTGTGAGG 630
 DB 1273 GGGAGCCCGATATTTTCCACATCTCCAGCCGGAATGACCAAGACTTGTGAGG 1332
 QY 631 GATGGCATGTCGCGTACGCTTCCATGATGATGATGATGATGATGATGATGATGAT 690
 DB 1333 TGAGCATGTCGCGTACGCTTCCATGATGATGATGATGATGATGATGATGATGAT 1392
 QY 691 TCGACACCTTACCAATATCTGATGATGATGATGATGATGATGATGATGATGAT 750
 DB 1393 TCGACACCTTACCAATATCTGATGATGATGATGATGATGATGATGATGATGAT 1452
 QY 751 GCGAAGCTTGTCTACTTGTGACCTTCAAGCTGAGAG--TGTGTCTTCAACAGGCTGC 807
 DB 1453 GCGGAGCTTGTCTACTTGTGACCTTCAAGCTGAGAG--TGTGTCTTCAACAGGCTGC 1512
 QY 808 ACAACCAAGAGAGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 867
 DB 1513 TCGCTGTGAGACCTCTTCTACCTTCAAGCTGATGATGATGATGATGATGATGATGAT 1572
 QY 868 TTCAACTCTTATGAGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 927
 DB 1573 TTCAACTCTTATGAGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1632

QY 928 CTGAGGCTGCTACCTGCTGCAAGGCGGGGCGCTTGGAGGAGCCCTTGAGTCCAT 987
DB 1633 CTGAGGCTGCTACCTGCTGCAAGGCGGGGCGCTTGGAGGAGCCCTTGAGTCCAT 1692
QY 988 CCAAGTATGATCCTGCTGCTTACCGCTTCTCCAGGCGCAGAGCGGAAATATCC 1047
DB 1693 AGCGAGAGCAGTATGCTTGGCTTCTCCAGGCGCAGAGCGGAAATATCC 1752
QY 1048 CTGATGAGTGGCGGCTGCTGCTTCTGATGAGGAGTAAATGAGCGGATTAAGAG 1107
DB 1753 CCGATGAGTGGCGGCTGCTGCTTCTGATGAGGAGTAAATGAGCGGATTAAGAG 1812
QY 1108 CGGCTGAGTGGCTGCTGCTTACCGGCGGAGGCGGAGCTGAGCTGCTGCTGAGGAG 1167
DB 1813 CGGCTGAGTGGCTGCTGCTTACCGGCGGAGGCGGAGCTGAGCTGCTGCTGAGGAG 1872
QY 1168 GACATCCCTGCGAGAGTGGCTTCTGATGAGGAGTAAATGAGCGGATTAAGAG 1227
DB 1873 GAGCTGAGTGGCGGCTGCTGCTTCTGATGAGGAGTAAATGAGCGGATTAAGAG 1932
QY 1228 AATGCTCCCTGCGAGAGTGGCTTCTGATGAGGAGTAAATGAGCGGATTAAGAG 1287
DB 1933 AATGCTCCCTGCGAGAGTGGCTTCTGATGAGGAGTAAATGAGCGGATTAAGAG 1992
QY 1288 GAGCGATGAGTGGCTGCTGCTTCTGATGAGGAGTAAATGAGCGGATTAAGAG 1347
DB 1993 GAGCGATGAGTGGCTGCTGCTTCTGATGAGGAGTAAATGAGCGGATTAAGAG 2052
QY 1348 ACCAAGTGGCGAGTGGAGTGGAGTGGAGTGGAGTGGAGTGGAGTGGAGTGGAG 1407
DB 2053 ACTAAGTGGCGAGTGGAGTGGAGTGGAGTGGAGTGGAGTGGAGTGGAGTGGAG 2112
QY 1408 TATGAGAGTGGAGTGGAGTGGAGTGGAGTGGAGTGGAGTGGAGTGGAGTGGAG 1464
DB 2113 TATGAGAGTGGAGTGGAGTGGAGTGGAGTGGAGTGGAGTGGAGTGGAGTGGAG 2172
QY 1465 AAGGAGTGGAGTGGAGTGGAGTGGAGTGGAGTGGAGTGGAGTGGAGTGGAGTGGAG 1524
DB 2173 AATGATGAGTGGAGTGGAGTGGAGTGGAGTGGAGTGGAGTGGAGTGGAGTGGAG 2232
QY 1525 TCTGATGAGTGGAGTGGAGTGGAGTGGAGTGGAGTGGAGTGGAGTGGAGTGGAG 1584
DB 2233 TCTGATGAGTGGAGTGGAGTGGAGTGGAGTGGAGTGGAGTGGAGTGGAGTGGAG 2292
QY 1585 TGGATGAGTGGAGTGGAGTGGAGTGGAGTGGAGTGGAGTGGAGTGGAGTGGAGTGGAG 1644
DB 2293 TGGATGAGTGGAGTGGAGTGGAGTGGAGTGGAGTGGAGTGGAGTGGAGTGGAGTGGAG 2352
QY 1645 CGAGGTTGGCTGCGAGATGAGCAGTGGTGGCTGAGCGGCTTCCATCCAGATATTC 1704
DB 2353 AATGAGTGGCTGCGAGATGAGCAGTGGTGGCTGAGCGGCTTCCATCCAGATATTC 2412
QY 1705 TCCGCTCTGAGTAAAGTGGCTGGCTGGCTGGAGTAAAGTGGCTGGAGTAAAGTGGAG 1764
DB 2413 TCAATATGAGCAGCAGCGGCTGGCTGGCTGGAGTAAAGTGGCTGGAGTAAAGTGGAG 2472
QY 1765 GCTGCGCTGAGTAAAGTGGCTGGCTGGCTGGAGTAAAGTGGCTGGAGTAAAGTGGAG 1824
DB 2473 GCGGATGAGTAAAGTGGCTGGCTGGCTGGAGTAAAGTGGCTGGAGTAAAGTGGAG 2532
QY 1825 CAGATGAGTAAAGTGGCTGGCTGGCTGGAGTAAAGTGGCTGGAGTAAAGTGGAGTGGAG 1884
DB 2533 CAGATGAGTAAAGTGGCTGGCTGGCTGGAGTAAAGTGGCTGGAGTAAAGTGGAGTGGAG 2589
QY 1885 GAGCAGATGAGTAAAGTGGCTGGCTGGCTGGAGTAAAGTGGCTGGAGTAAAGTGGAGTGGAG 1944
DB 2590 GAGCAGATGAGTAAAGTGGCTGGCTGGCTGGAGTAAAGTGGCTGGAGTAAAGTGGAGTGGAG 2649
QY 1945 ACCAGCTTGGCTTCTGATGAGGAGTGGAGTGGAGTGGAGTGGAGTGGAGTGGAGTGGAG 2004
DB 2650 ACCAGCTTGGCTTCTGATGAGGAGTGGAGTGGAGTGGAGTGGAGTGGAGTGGAGTGGAGTGGAG 2709

QY 2005 CCAATGAGTGGCTGAGTGGAGTGGAGTGGAGTGGAGTGGAGTGGAGTGGAGTGGAGTGGAG 2064
DB 2710 GCTTCCGCTGAGTGGAGTGGAGTGGAGTGGAGTGGAGTGGAGTGGAGTGGAGTGGAGTGGAG 2769
QY 2065 TCTTCCAGAGAGCGGAGTGGAGTGGAGTGGAGTGGAGTGGAGTGGAGTGGAGTGGAGTGGAG 2124
DB 2770 TCTTCCAGAGAGCGGAGTGGAGTGGAGTGGAGTGGAGTGGAGTGGAGTGGAGTGGAGTGGAG 2829
QY 2125 AAGATGCTGAGTGGAGTGGAGTGGAGTGGAGTGGAGTGGAGTGGAGTGGAGTGGAGTGGAG 2184
DB 2830 GAGATGCTGAGTGGAGTGGAGTGGAGTGGAGTGGAGTGGAGTGGAGTGGAGTGGAGTGGAG 2889
QY 2185 CAGCGGAGTGGAGTGGAGTGGAGTGGAGTGGAGTGGAGTGGAGTGGAGTGGAGTGGAGTGGAG 2244
DB 2890 CAGCGGAGTGGAGTGGAGTGGAGTGGAGTGGAGTGGAGTGGAGTGGAGTGGAGTGGAGTGGAG 2949
QY 2245 CAGGAGTGGAGTGGAGTGGAGTGGAGTGGAGTGGAGTGGAGTGGAGTGGAGTGGAGTGGAG 2304
DB 2950 CAGGAGTGGAGTGGAGTGGAGTGGAGTGGAGTGGAGTGGAGTGGAGTGGAGTGGAGTGGAG 3009
QY 2305 TATGAGGAGTGGAGTGGAGTGGAGTGGAGTGGAGTGGAGTGGAGTGGAGTGGAGTGGAGTGGAG 2364
DB 3010 TATGAGGAGTGGAGTGGAGTGGAGTGGAGTGGAGTGGAGTGGAGTGGAGTGGAGTGGAGTGGAG 3069
QY 2365 TTCAATGAGTGGAGTGGAGTGGAGTGGAGTGGAGTGGAGTGGAGTGGAGTGGAGTGGAGTGGAG 2424
DB 3070 TTCAATGAGTGGAGTGGAGTGGAGTGGAGTGGAGTGGAGTGGAGTGGAGTGGAGTGGAGTGGAG 3129
QY 2425 GAGAGTGGAGTGGAGTGGAGTGGAGTGGAGTGGAGTGGAGTGGAGTGGAGTGGAGTGGAGTGGAG 2484
DB 3130 GAGAGTGGAGTGGAGTGGAGTGGAGTGGAGTGGAGTGGAGTGGAGTGGAGTGGAGTGGAGTGGAG 3189
QY 2485 CAGAGTGGAGTGGAGTGGAGTGGAGTGGAGTGGAGTGGAGTGGAGTGGAGTGGAGTGGAGTGGAG 2544
DB 3190 GAGAGTGGAGTGGAGTGGAGTGGAGTGGAGTGGAGTGGAGTGGAGTGGAGTGGAGTGGAGTGGAG 3249
QY 2545 TCTGAGTGGAGTGGAGTGGAGTGGAGTGGAGTGGAGTGGAGTGGAGTGGAGTGGAGTGGAGTGGAG 2604
DB 3250 TCTGAGTGGAGTGGAGTGGAGTGGAGTGGAGTGGAGTGGAGTGGAGTGGAGTGGAGTGGAGTGGAG 3309
QY 2605 CCGCGAGAGGAGGAGTGGAGTGGAGTGGAGTGGAGTGGAGTGGAGTGGAGTGGAGTGGAGTGGAG 2664
DB 3310 CCGCGAGAGGAGGAGTGGAGTGGAGTGGAGTGGAGTGGAGTGGAGTGGAGTGGAGTGGAGTGGAG 3369
QY 2665 GAGATGCGCTCCATGAGTGGAGTGGAGTGGAGTGGAGTGGAGTGGAGTGGAGTGGAGTGGAGTGGAG 2724
DB 3370 GAGATGCGCTCCATGAGTGGAGTGGAGTGGAGTGGAGTGGAGTGGAGTGGAGTGGAGTGGAGTGGAG 3429
QY 2725 ATCCGAGAGTGGAGTGGAGTGGAGTGGAGTGGAGTGGAGTGGAGTGGAGTGGAGTGGAGTGGAGTGGAG 2784
DB 3430 ATCCGAGAGTGGAGTGGAGTGGAGTGGAGTGGAGTGGAGTGGAGTGGAGTGGAGTGGAGTGGAGTGGAG 3489
QY 2785 TTGAGTGGAGTGGAGTGGAGTGGAGTGGAGTGGAGTGGAGTGGAGTGGAGTGGAGTGGAGTGGAGTGGAG 2844
DB 3490 CCAATGAGTGGAGTGGAGTGGAGTGGAGTGGAGTGGAGTGGAGTGGAGTGGAGTGGAGTGGAGTGGAGTGGAG 3549
QY 2845 TATTAATGAGTGGAGTGGAGTGGAGTGGAGTGGAGTGGAGTGGAGTGGAGTGGAGTGGAGTGGAGTGGAGTGGAG 2904
DB 3550 TATTAATGAGTGGAGTGGAGTGGAGTGGAGTGGAGTGGAGTGGAGTGGAGTGGAGTGGAGTGGAGTGGAGTGGAG 3009
QY 2905 GAGAGTGGAGTGGAGTGGAGTGGAGTGGAGTGGAGTGGAGTGGAGTGGAGTGGAGTGGAGTGGAGTGGAGTGGAG 2964
DB 3610 GAGAGTGGAGTGGAGTGGAGTGGAGTGGAGTGGAGTGGAGTGGAGTGGAGTGGAGTGGAGTGGAGTGGAGTGGAG 3069
QY 2965 TTGAGTGGAGTGGAGTGGAGTGGAGTGGAGTGGAGTGGAGTGGAGTGGAGTGGAGTGGAGTGGAGTGGAGTGGAG 3021
DB 3670 CCGCGAGAGGAGGAGTGGAGTGGAGTGGAGTGGAGTGGAGTGGAGTGGAGTGGAGTGGAGTGGAGTGGAG 3081
QY 3022 ACCAGCTTGGCTTCTGATGAGGAGTGGAGTGGAGTGGAGTGGAGTGGAGTGGAGTGGAGTGGAGTGGAG 3081
DB 3730 CCGCGAGAGGAGGAGTGGAGTGGAGTGGAGTGGAGTGGAGTGGAGTGGAGTGGAGTGGAGTGGAGTGGAG 3789
QY 3082 ATCCAGAGAGTGGCTTCTGATGAGGAGTGGAGTGGAGTGGAGTGGAGTGGAGTGGAGTGGAGTGGAGTGGAG 3141

Db 3790 GTGATAGCAACCTGAGTTGAGTACATAGATGACCTCGGGCTCCAGCGCATGAGCCA 3849
Qy 3142 GAATGAGATGTTGAGTGAAGAACACACCCATGGCGGTATGGGGAGACCCACCTGAGACCTC 3201
Db 3850 GAGTGGAGCATTTGCGAGTGGCCACACACCTGACATCACAGGCTTCAACCTGGATGTC 3909
Qy 3202 ATACGAACCCCCAGATCTGTCGCAAGCATGAGAGGAAGAGCAATCAATATCTGTGAG 3261
Db 3910 ATTCGAGAGCCCAAGATCCGATCAAAATTCATGAGCAAAAGATCTGTCAATGTGTAA 3969
Qy 3262 GTTCTGAGCTACTGAGATGATGATCTGAGGCGCCCGCTCGCTGAGGCTCTGACAC 3321
Db 3970 GTTGTGAGACACACACCTCACTGCTGAGCACTCTGTGACACGATACCGGCTT 4029
Qy 3322 CAGTGAAGCTGACGAGAGGCGAGAGGTTTGGCTTCACTCTGACACACGTCCTGCTC 3381
Db 4030 GGCTGAGACATGTGGAAGCGCCAGATGAGTTGATTTGTCTTAACTATGTCATATCC 4089
Qy 3382 CTGCTCATCTCTCAACAGACCACTTCACTATCTATCCAAACCGAGTGTGAGGCTT 3441
Db 4090 TTGCTAATTACACAGACACCAAGTTATCTACTACCCCAACCCGACCTTGAATGCTT 4149
Qy 3442 GGTTCCTCAGGAATCTCTGAGCTCAGGCTGAGCGCCCATGATCTCTAAAGGCAAGAC 3501
Db 4150 AACCTTACTGAGATCTTGGATCAAAAGCCAGAGATGCGCCATCTTGAAGGCAAAAC 4209
Qy 3502 CTGATCTCCGCTGTGGCTGGGGGCAAGTGAAGCTGATCACTGTGCTGTGAGGAG 3561
Db 4210 CTCTGCTCTGCTCTGAGAG--GCGCAACTCACTGAGCTCATCGAGAG 4266
Qy 3562 AAGCGGTGACCGGTGACCGGTGATGATGTCAGTGTGCTGAGTCCGAGTCCCAACCTCAG 3621
Db 4267 ACCCTTGTGCTGTACCGGTATCTGAGACCCAGTCTTCTGCGAGGCTCCCAACCTCAG 4326
Qy 3622 GGCAGGACCAAGTATGAGCCCGTGTGAGGAGTGAATCTCCCGGAGATGTGTAC 3681
Db 4327 GGGCAGACCAAGTATGATGTCAAGTGGGCGGATGTGTCTCGCTGCTGAGT 4386
Qy 3682 ATTGCGCGGACAGCCCGCTCAGGCTGCGCCATGCTGACATGAGAGTGTGAGGCT 3741
Db 4387 GTCATCTCAGACAGTGTGTGACCTTCCAGCCATCTCCAGTCCGCGGCGGCGAGC 4446
Qy 3742 CTCTCATCATTTTATGCTGAGCGGTCTCATTTGCTTAAAGCAATCCCGCAAGT 3801
Db 4447 CTCTCTCATCATGATGATCATGATCTCTCATTTGCTTAAAGCGCAAGTCTCGAAGAAAT 4506
Qy 3802 GACCTCAAGCTGAAGCGGCTGATGAGTGAAGCAACTGAGTCCCGTGTGAGGCTG 3861
Db 4507 GACCTCACTCTCAAGCGGCTGCAATGAGATGAGCAATCTGAGTCCCGTGTGAGGCTG 4566
Qy 3862 GAGTGAAGAGAGCTTGTGCGAGGTGAGACGACATCCATGAGCTGAGCATGACCTG 3921
Db 4567 GAGTGAAGAGAGCTTGTGCGAGGTGAGACGACATCCATGAGCTGAGCATGACCTG 4626
Qy 3922 GATGAGCCGAGATCTCTCTCTGAGTATGAACTTACACATGCGGCTGTCTTCCA 3981
Db 4627 GACCGCTCAGAACTCTTACCTGAGATGATGATCTGAGGAGTCCGCGCTGAGGAGCT 4041
Qy 3982 GGAATTGAAGACACCGCTGCTCTCGGAGCTTGAAGTCCCGGCTGAGCGGAGGAGCT 4041
Db 4687 GGCATGAGAGACACCGCTGCTCTCGGAGCTTGAAGTCCCGGCTGAGCGGAGGAGCT 4746
Qy 4042 GTGAGAGAGAGCTTGAAGCTTGTGCGCAGCTCATCAACACAGAGTGTCTGCTGCTC 4101
Db 4747 GTGAGAGAGAGCTTGAAGCTTGTGCGCAGCTCATCAACACAGAGTGTCTGCTGAGC 4806
Qy 4102 TTCAATCCGAGCTTGAAGCTTGTGCGCAGCTTCAACAGCGAGCTGAGCAAGCTGAGC 4161
Db 4807 TTCAATCCGAGCTTGAAGCTTGTGCGCAGCTTCAACAGCGAGCTGAGCAAGCTGAGC 4866
Qy 4162 TCATCATCATGACCTGTGTCAGAGCAAGCTGAGTACGCTGATGTGTGAAGAG 4221

Db 4867 TCGCTCATATGACCGGCTGAGAGGCGCCTGGAATATGCACTGATGTCTTCAAGAG 4926
Qy 4222 CTGCTGCGGACCTCTTATGACAAAGCTGAGAGAGCAAGACCACTTAAGCTGTCTC 4281
Db 4927 CTGCTCTGACCTCATATGATTAAGACTGAGAGAGCAAGACCACTTAAGCTGTCTC 4986
Qy 4282 AGAGAGATGAGTCAATGCTGAGAGAGTCTGAGAGCAATTTGATTTACTTCTCTCTAC 4341
Db 4987 CGAGAGACAGAGTCTGTGCTTGAAGAGTCTGAGCAATTTGATTTGCTCTCTCTGAC 5046
Qy 4342 AAGTTCCTAAGAGTGTGCTGAGAGGCGCTCTTCTCTCTCTGATCTGAGCATGAGAG 4401
Db 5047 AAGTTCCTAAGAGTGTGCTGAGAGGCGCTCTTCTCTCTGATCTGAGCATGAGAG 5106
Qy 4402 CAGATGAGAGAGGCGCTTGAAGCGCTCAACCGGCGAGGCGCTCTCTCTGAGAG 4461
Db 5107 CAGATGAGAGAGGCGCTTGAAGCGCTCAACCGGCGAGGCGCTCTCTCTGAGAG 5166
Qy 4462 GACAGCTCATCGCGCAGAGATGATGATCAAAACCTTGTCTGAGCTGTGAGCTCA 4521
Db 5167 GACAGCTCATCGCGCAGAGATGATGATCAAAACCTTGTCTGAGCTGTGAGCTCA 5226
Qy 4522 GACATGCGACAGCGCGGAGGCTCCAGTAAAGATCCCTCACTGTGACCATGACCTGAG 4581
Db 5227 GACATGCGACAGCGCGGAGGCTCCAGTAAAGATCCCTCACTGTGACCATGACCTGAG 5286
Qy 4582 GTCAAGAGAGAGTGTGATGATGATCTTCAAGATGTGCTTCTCCACCGGCGCA 4641
Db 5287 GTCAAGAGAGAGTGTGATGATGATCTTCAAGATGTGCTTCTCCACCGGCGCA 5346
Qy 4642 GTGCAAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 4701
Db 5347 GCAATGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 5406
Qy 4702 GAAGATCATCAACCAAGATTTGAATTTGAATTTGAATTTGAATTTGAATTTGAATTTGA 4761
Db 5407 GAGGATCATCAACCAAGATTTGAATTTGAATTTGAATTTGAATTTGAATTTGAATTTGA 5466
Qy 4762 CAGGTGCAAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 4821
Db 5467 CAGGTGCAAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 5526
Qy 4822 GTGAACATCTCAAGCGCTTCCAGAGCTCAGAGATTAATTTGAATTTGAATTTGAATTTGA 4881
Db 5527 CTGCTCTGCGAGATCTTCCGAGCTCATGAGATTAATTTGAATTTGAATTTGAATTTGA 5586
Qy 4882 ACCGAGAGCGCGGACAGCTTCCGAGCTCATGAGATTAATTTGAATTTGAATTTGAATTTGA 4941
Db 5587 ACCGAGAGCGCGGACAGCTTCCGAGCTCATGAGATTAATTTGAATTTGAATTTGAATTTGA 5646
Qy 4942 GAGATCAAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 5001
Db 5647 GAGATCAAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 5706
Qy 5002 CGGAGAGAGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 5061
Db 5707 CGGAGAGAGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 5766
Qy 5062 CTGAGAGATTTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 5121
Db 5767 CTGAGAGATTTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 5826
Qy 5122 GCGCTGCGCTGCGCTCAAGTATGATGATGATGATGATGATGATGATGATGATGATGATGAT 5181
Db 5827 GCTCTGCGCTGCGCTCAAGTATGATGATGATGATGATGATGATGATGATGATGATGATGAT 5886
Qy 5182 GGCATTCATGACCGGACCTCGCATCTGAGAGAGCAATTTGATTTGATTTGATTTGATTTGAT 5241
Db 5887 AGCATCATGACAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 5946
Qy 5242 TGGGTCAATGATCAAGAGCTGATGATGATGATGATGATGATGATGATGATGATGATGAT 5301
Db 5947 TGGGTCAATGATCAAGAGCTGATGATGATGATGATGATGATGATGATGATGATGATGAT 6006

QY 1048 CTGATAGAGTGGGCGGCTGGATCTTCACTTGAAGCAGATTAATGACCGCATTAAGAG 1107
DB 1321 CCGATGAGTCTGCGCTGTGTGCTTCCCTATCCGGGCGCATCAATGGAGATCAAGAG 1380
QY 1108 CCGCTGAGTCTTGTATACCGGGGCGAGGCGACGCTGACCTGGCTGCTCAAGGTGAAG 1167
DB 1381 CCGCTGAGTCTTGTATACCGGGGCGAGGCGACGCTGAGCTCAACTGGCTGCTGAGGAG 1440
QY 1168 GACATCCCTGAGAGAGTGGCTCTTAACAATTGACATTAATCTTGTGGCTGAGCATG 1227
DB 1441 GACGTGAGTGAACCAAGCGCGCTGTCCCATGATGATTAATCTTGTGAGCTGAGCATC 1500
QY 1228 AATGCTCCCTGGAGAGTGGCGACATGGTGGTGAATTCGGATCTCAAGAGAGAG 1287
DB 1501 AACGAGCCCTGGAGAGGCTCAACTCGAGTGAAGGCGCTGACCTGTACACACAGAGAG 1560
QY 1288 GACCGATGACGTTCTGATGACATATGTCTTAACAAGAACATCTCTGGCTTGTGGG 1347
DB 1561 GACCGATGACGTTCTGATGACATATGTCTTAACAAGAACATCTCTGGCTTGTGGG 1620
QY 1348 ACCAAAAGTGGCAAGCTGAAGAGATCCGGGTGATGAGACCGAGGCGAACCGCTCCAG 1407
DB 1621 ACTTAAGTGGCAAGCTGAAGAGATCCGGGTGATGAGACCGAGGCGAACCGCTCCAG 1680
QY 1408 TATGAGACGCTGAGAGTGT--GACCCCGGCTCAATCTCTGGGATATGGCTTCTCC 1464
DB 1681 TACGATGATGCTCTGTGCTCAAGAGCGAAGCCCATCTCCGGGACATGGCTTCTCC 1740
QY 1465 AAGGACCAAGAGCACTGATCATGTCAGAGAGGAGCTCAACAAGTCCCTGTGAG 1524
DB 1741 ATGATCAAGCTGATCTGATGATGATGATGATGATGATGATGATGATGATGATGAT 1800
QY 1525 TCGTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1584
DB 1801 TCGTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1860
QY 1585 TGTGTGTGCTGACAAACGTCGACCCGGAAGAGCGGTGTAGACGCTCAAGAGAGCC 1644
DB 1861 TGTGTGTGCTGACAAACGTCGACCCGGAAGAGCGGTGTAGACGCTCAAGAGAGCC 1920
QY 1645 CCGAGGTTGCTCGAGAGTGAAGCAGTGTGCGGCTGACCGGCTCAATATTC 1704
DB 1921 AATCATTTGCTCGAGAGTGAAGCAGTGTGCGGCTGACCGGCTCAATATTC 1980
QY 1705 TCGTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1764
DB 1981 TCGTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2040
QY 1765 GCTGCGTGAAGTGAAGCTTTGAGAGCTGTCAAGATGATGATGATGATGATGATGAT 1824
DB 2041 GCGGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2100
QY 1825 CAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1884
DB 2101 CAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2157
QY 1885 GACGACCATGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1944
DB 2158 GACGACCATGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2217
QY 1945 ACCAGCTTGTCTTCAATGAGAGCTGACATTAATGATGATGATGATGATGATGATGAT 2004
DB 2218 ACCAGCTTGTCTTCAATGAGAGCTGACATTAATGATGATGATGATGATGATGATGAT 2277
QY 2005 CCAATACCGTGAAGTGTGATTAATCGGAGATGATGATGATGATGATGATGATGATGAT 2064
DB 2278 GCTTTCGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2337
QY 2065 TCGTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2124
DB 2338 TCGTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2397
QY 2125 AATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2184

DB 2398 GAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2457
QY 2185 CAGCCGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2244
DB 2458 CAGCCGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2517
QY 2245 CAGCCGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2304
DB 2518 CAGCCGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2577
QY 2305 TATGAGGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2364
DB 2578 TATGAGGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2637
QY 2365 TATGAGGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2424
DB 2638 TATGAGGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2697
QY 2425 GAGAGTGGGCTGTGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2484
DB 2698 GAGAGTGGGCTGTGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2757
QY 2485 CAGGCGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2544
DB 2758 GAGCGAGTGAAGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2817
QY 2545 TGTGTGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2604
DB 2818 TGTGTGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2877
QY 2605 CCGCGGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 2664
DB 2878 CCGCGGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 2937
QY 2665 GACATGCTGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2724
DB 2938 GACATGCTGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2997
QY 2725 AATCCCTGAGAGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2784
DB 2998 AATCCCTGAGAGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 3057
QY 2785 TGTGTGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2844
DB 3058 CAGATGCTGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 3117
QY 2845 TATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2904
DB 3118 TATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 3177
QY 2905 GAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 2964
DB 3178 GAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 3237
QY 2965 TTTGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 3021
DB 3238 CTTGGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 3297
QY 3022 ACCAGATCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 3081
DB 3298 ACCAGATCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 3357
QY 3082 ATCCAGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 3141
DB 3358 GTGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 3417
QY 3142 GATGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 3201
DB 3418 GATGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 3477
QY 3202 ATGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 3261

Db	3478	ATTGAGAGGCAAGATCGAGTCAAAATTCAATGACAAAGAAATCTGATCAATGATGTGTAA	3537
Oy	3562	GTTCTBAAGCATATGAGATGACCTGTGATAGAGGCGCGCCCTGCGTCTGGGTCTCTAAGCAC	3521
Db	3538	GTTGTGAACAAACACACCTTCACTGCTGGACACCTCTTGTGACACAGCATACCGCCCT	3597
Oy	3522	CAGTCAGACTGACCGAGAGGCCGAGAGATTGAGTTCATCTGTGACAAAGTCCAATGCC	3581
Db	3598	GGCCTGACACTGTGAGAGGCCCAATGATTTGTGATTTGTCTTAAACAATGTCCAAATCC	3657
Oy	3582	CTGCTCATCTCTCAACAAGACCAACTTCACTCACTATCCCAACCGGATGTTAGAGCTTT	3441
Db	3658	TTTGCTAATTACAAACAGACCAAGTTTATTACTAACCCCAACCCGACCTTTGAACTGCTT	3577
Oy	3442	GGTCCCTCAGAGATTCCTGGAGCTCAAGCCTGACAGCGCCAAATCTCTAAAGGCAAGAC	3501
Db	3718	AGCCTTCTGAGATCTTTGATTAAGAGAGAGATGGCCATCACTTGAAAGGCAAAAC	3777
Oy	3502	CTGATCCCGCTGTGGCTGGGGGCAACGTAAAGTGAATCACTGTGCTGTGGGGAG	3561
Db	3778	CTGTGCTCTCTGCTCTGTGAGAG--GGCCAAACTCAACTACATGTGCTCATCGGAAG	3854
Oy	3562	AAGCCGTCACCGTACCGGTGCAATATCCAGCTGCTGTGCGAGTCCCCCAACTCATC	3621
Db	3835	ACCCTTGTGTCTGTCAACCTATCTAGAACCCAGTTCTGTGCGAGCTCCCACTTCAAC	3894
Oy	3622	GGCAGGCAACAAGTATGAGCCCGTGTCCGTGGCATGAGATCTCCCGGGAGTGTATAC	3661
Db	3895	GGGAGACACAAGTCTTGGTTTACAGTGGGCGGATGGTGTCTGGCTTGGTGGT	3954
Oy	3682	ATTGCCCGGACAGCCCGCTGACCTGTGCCCGCATCTGCATGCAATGCTGGCGGC	3741
Db	3955	GTCATCTCAAGACAGTTGTGTGACCTGTGCACCATCTGCAGATCCGCGCGGGCAGC	4014
Oy	3742	CTCTCATATTTTATCGTGGCCGCTGTCAATGCTCTATTAAGCAAGTCCCGGAAAGT	3801
Db	4015	CTCTCTCTATATGTGTATATGTCTTCTTGTGCTTCAAGGCAAGTCTGAGAAAT	4074
Oy	3802	GACCTCACGCTGAAGCGGCTGCAGATGACAAATGACAACTGGAATGCCGTGGCCCTG	3861
Db	4075	GACCTCACTCTCAAGGGCTGCAAATGACAAATGACAACTGGAATGCCGTGGCCCTG	4134
Oy	3862	GAGTGAAGGAAGCTTTTACCAGGTGCAGACGGACATCCATAGGTGACCAATGACTG	3921
Db	4135	GAGTGAAGGAAGCTTTTGTGAGCTTCAGACGGATATCAATGATGTGACATGACTG	4194
Oy	3922	GATGAGACCGGATTCCTGCTGTCGATCATATGAATCTTACCATATGGGGGTGCTTCCA	3961
Db	4195	GACCGCTCAGGAATCCCTTACCTGGAATGATCTACACGTATGCGAGTCTGTGTTCCG	4254
Oy	3982	GGAATGGAAGACCACTGTGCTCTCCGGACCTTGAAGTCCGGGCTTACCGGACGAGCT	4041
Db	4255	GGCATTCAGAGACACCCGCTCTGTGGGAGGTGAGAGGTATAMAGAAACGGGACGAGAC	4314
Oy	4042	GTCGAGAAAGGCTGAGAGCTTTCGCGCAGCTCATCAACAACAAGGTGTTCTGTGCTCC	4101
Db	4315	GTCGAGAGAGCCCTGAAGCTTTTGTCCAGCTCATCAACAACAAGGTGTTCTGTGACAC	4374
Oy	4102	TTTATTCGACGCTTGAAGTCCAGGTGATCTTCTCATATGGGCAACGATGGCAACGTGCC	4161
Db	4375	TTTATTCGACGCTTGAAGTCCAGGTGATCTTCTCATATGGGCAACGATGGCAACGTGCC	4434
Oy	4162	TCACTATCATGACGCTGTGTCGAGACAAAGCTGGAAGTACGCACTGATGTGCTGAAGAG	4221
Db	4435	TGCGTATCATGACCGGCTCTGACGGCGCGCTGGAATATCCATGATGTCTCTCAAGAG	4494
Oy	4222	CTGCTGGCCGACTCATATGACAAGAACTGAGAGCAAGAACCAACCTTAAGCTGTGCTGC	4281
Db	4495	CTGTCTCTGAACTCATGTATAGAACTGGAGAAACAAGAACCAACCCCAAGTGTGATATC	4554
Oy	4282	AGAGAGCATGAGTCAATGTGCTGAGAAATGCTGACCAATTTGTTACTTTCTCTCTTAC	4341
Db	4555	CGAGGACACAGTCTGTGGCTGAAAGATGCTGACCAATTTGTGGCTTCTCTCTGACAC	4614

OY	4342	AAAGTTCCTCAAGAAAGTGGAGTGGCGGGGAGACCCCTCTCTCCCTGGTCTCGTGGCCATCAAGAG	44910
Db	4615	AAAGTTCCTCAAGAAAGTGGAGTGGCGGGGAGACCCCTCTCTATGTATATCTGTGCCATCAAGAG	46744
OY	4402	CAGATGGAGAAAGGCCCCATTTGACGCCATCAACGAGCGAAGGCCCGGTACTCTCTTGAAGCGAG	4461
Db	4675	CAGATGGAAAGAGGCCCATTTGAGCCATCAACGAGCGAAGGCCCGGTACTCTCTTGAAGCGAG	47344
OY	4462	GACAACTCATCCGCGACAGATTGACTAAGAAACCCTGGTCTCGAGCTGTCAAGCCCA	45311
Db	4735	GACAACTCATCCGCGACAGATTGAGTGAACAGACCCTGATCTTGAACTGGGTCAACCT	47344
OY	4522	GACAAATGCCAAAGCCCCGAGAGTCCCAAGTAAATCTCTGAATGAGACCAATCACTCACTCA	45811
Db	4795	GACAAAGAACAGTCCACAGATCCCAAGTAAAGGTGTAACTGTGACACCAATCAACAG	4854
OY	4582	GTCCAGAGAAAGATTCTGAGATGCCATCTTCAAGATGTGCCTTGCTCCACCGGCCCAA	46411
Db	4855	GTCCAGAGAAAGATTCTTGATGCCGTGTAAAGAAATGTGCCATATCCAGCGGCGAG	49144
OY	4642	GCTGCAGATATGAACTCTGGAATGGCGCAACAAGAAAGTGGGACCAAGATGATCTTGAGAGAT	47010
Db	4915	GCAATGACATGGAATTTGAGATGGCGCAAGGCCGAGATGGCCGGGTGGTGTGAAAT	49744
OY	4702	GAAACATATCAACCAAGATTGAGATGATTGAAAGCACTGAACAACCTGGCCCACTAC	4761
Db	4975	GAGGAAATCAACCAAGATTGAGAGTGAATGGAACGGGTCAACCACTGATGTGATAT	50344
OY	4762	CAGGTGCCAGATGATCTCTGTGGTGGCATTTAGTGTCTCAACAGAGTGAACGCTTAAACCA	48211
Db	5035	CAGGTGTCAAGAGGTGGTGTGTGGTCTCTGTGCCCAACAAGCCCTCTCAACAATC	50944
OY	4822	GTGAACAATCACTCAACCGTCTCCAGGACCTCAGAAAGTAAATGAAAGAACTGATCCGAGAT	4881
Db	5095	CTGTGCTCTGCCAGATCTCCCGACGTCCATCAGAGATATGATCTCTCTCAAGAT	51544
OY	4882	ACGGGACGCCCCGACAGCTTCGGTCAACGACCTATGATCACTCTGACCTGGAAGT	49411
Db	5155	ACGGGAGCCCCGACAGCTGGGATCCCGGGCCCCATGATCAACCCAAACCTGGAAGT	52144
OY	4942	GGAGTCAAGATTTGGCACTTAGTAAAGAACACAGACACGAGACACAGAGAGGGGAG	50010
Db	5215	GGAGTCAAGTGTGCACTTAGTAAAGAACATGACACAGTGAACAGAGAGGGGTAC	52744
OY	5002	CGGGGAGAGAAATGGTGTGAAATCTCACTCAACCGGACTCTGGGCACTTAAAGGGCA	5061
Db	5275	CGGGGAGAGAAATGGTGTGCAATCTCACTCAACCGGCTACTGGCACACAGGCGAAC	53344
OY	5062	CTGCAGAAAGTTGTGATGACCTCTTTGAGACCATCTTCAACCGGACACCGTGGCTT	5121
Db	5335	CTGCAGAAAGTTGTGAGACCTTCTTTGAGACCTTGTACGACCTGTGACACCGGGGAGC	53944
OY	5122	GCCCGGCCCTGGGCATCAAGTATGATTTGACTTCTGAGATGAGACGGCTGTAACAT	51811
Db	5395	GCTCTCCCTGGCCATCAAGTATGATTTGATTTCTGATGAGAGGACAGACGAC	54544
OY	5182	GGGATTCATGACCGGCAAGTCCGCAATCTGGAAGAGAAATGGCTGGCCCTGGAGGTT	52411
Db	5455	AGGATTCATGACAGATGTGCGGACACCTGGAAGAAAGCACTGGCTCCTCTGGGCTTC	55144
OY	5242	TGGGTCAACATGATCAAGAACCCGCACTTTGTGTGACATCATTAAGAACGATCA	53010
Db	5515	TGGGTCAACATGATTAAGAACCCCAATTTGATTTGACATCAACAAGGCGAGTACG	55744
OY	5302	GACGCTGGCTCTCTGTGTGGGTCAACATTTATGAACTCTTGCTCCACGTCAGAGAC	53611
Db	5575	GACGCTGGCTCTCTGTGTGGGCGCAACCTTATGACATCTTTGATCAACGTCAGAGAC	56344
OY	5362	CGGCTGGGCAAGAACTCGCCCTCTCAACAAGCTCTGTATGCCAAGACATCCCGAGCTAC	54211
Db	5635	CGGCTGGGCAAGAACTCCCTCTCAACAAGCTCTCTATGCAAGGACATCTCCAGCTAC	56944

QY 5422 AAGATTGAGTGGAGGATTAATCTACATAGAGAGATCCAGCATCAGGACCAA 5481
 DB 5695 AAGAGCTGGGTGGAGATATCTACGACAGATGGCCAACTCCAGCATCAGTACAG 5754
 QY 5482 GACATGAACCGATACCTGGTGGAGCATGCCGATGACATGAATGAACCATCAG 5541
 DB 5755 GACATGATCCCTACCTCCGCGAGAGTCCCGCTGACGCGCTGAGTTCAACATGCTG 5814
 QY 5542 AGTCACTCTCAGAGATCTTCTCTATGTGGCAATACAGGAGAGATCTTGAACCT 5601
 DB 5815 AGTCCCTCAATGAGATCTACTCTATGTCAAGCAATATAGAGAGAGTCACTGAGGCC 5874
 QY 5602 CTGACACACATGATACCGCTGTGGAGAGCAAACTGCGCTACAACTGAGCAAGTCTA 5661
 DB 5875 CTAGAGCAGATGAGCAGGACGCGCGAGCGGCTGCTTATAGGTGAGAGCTCAT 5934
 QY 5662 ACCCTCATGAGCTTGAACAGCTGAAA 5687
 DB 5935 AATGCCATGTCCATTGAGAGCTGAGA 5960

RESULT 5

AB199504
 ID AB199504 standard; cDNA; 6730 BP.

AC AB199504;

DT 07-MAR-2002 (first entry)

DE Mouse ischaemic condition related cDNA sequence SEQ ID NO:479.

XX Mouse; ischaemia; compressive ischaemia; occlusive ischaemia;
 KM vasospastic ischaemia; ischaemic condition; ischaemic disease; ss.

XX Mus musculus.

XX W020018188-A2.

XX 22-NOV-2001.

XX 18-MAY-2001; 2001KO-JP004192.

XX 18-MAY-2000; 2000JP-00145977.

XX (UTNI-) UNIV NIHON SCHOOL JURIDICAL PERSON.

XX Ishikawa K, Asai S, Takahashi Y, Nagata T, Ishii Y;

XX WPI; 2002-034733/04.

XX P-PSDB; ABB57196.

XX Examining the ischemic condition (e.g. occlusive ischemia) by measuring
 PT expression levels of particular genes defined in the specification or by
 PT determining the expression profile of a gene group comprising these
 XX genes.

XX Claim 2; Page 1319-1331; 2690pp; English.

XX The present invention describes a method for examining ischemic
 CC conditions, comprising measuring the expression levels of particular
 CC genes (I) in a test sample or determining the expression profile of a
 CC gene group in the sample comprising genes selected from (I). The method
 CC is useful for examining the ischemic condition (e.g. compressive
 CC ischaemia, occlusive ischaemia or vasospastic ischaemia) by measuring the
 CC expression levels of particular genes (AB199202 to AB199912, encoding the
 CC protein sequences in ABB57020 to ABB57374) or by determining the
 CC expression profile of a gene group comprising these genes. The expression
 CC levels or expression profiles produced by these genes are used as an
 CC indicator when screening for ischaemic condition-improving drugs or
 CC therapeutics for ischaemic diseases. AB199913 and AB199914 represent PCR
 CC primers for a mouse ischaemic condition related sequence, which are used
 CC in the exemplification of the present invention

SQ Sequence 6730 BP; 1656 A; 1850 C; 1740 G; 1482 T; 0 U; 2 Other;
 Query Match 48.8%; Score 2778.4; DB 6; Length 6730;
 Best Local Similarity 69.3%; Pred. No. 0;
 Matches 3858; Conservative 0; Mismatches 1691; Indels 15; Gaps 5;

QY 137 AGCCGCGAGGGTTTCATATCCCTGTGTGTGATGAGAGACAGACATTTACTTGG 196
 DB 683 ATCTGATGTGACCTTTCACCATTTGACTGTATACCGAAGAACAGGGGCTGTGATGTG 742
 QY 1297 GGGCCGTCAATCGGATTTACAGCTCTCCAGCAGCTGAGAGTCTGTGACGATGAA 256
 DB 743 GGGCATGATCTGTCTTACAGATTTGACTGGCAACTGCACATCCAGTGGCTCACAAG 802
 QY 257 CAGGGCCGAGAGAGACACACCCCAAGTTTACCCACCCGCACTGTCCAGACTTGAATG 316
 DB 803 CAGGGCCGAGAGAGACACACAGGCTTGTACCCACCCCTCATTTGACAGCCCTGCAATG 862
 QY 317 AGCCCTGACACACACACACATGTCTACACAGATGTCTCTCATATAGCTACAAAGAGACA 376
 DB 863 AAGTCTTACACTCACCACACATGTCTACAACTATGATCATTTGACTACTGTGAAATC 922
 QY 377 GCGTATTTGCTGTGGAGAGCTGTACCAAGCATCTGCAAGCTGTGAGGAGGAGC 436
 DB 923 GCTGTCTGGCTGTGGAGAGCTCTACCAAGGCTGTGCAAGCTCTGTGCACTGATGACC 982
 QY 437 TCTTCAGCTGGGGAGAGCTTATCATTAAGAGAGACATCTGTGAGGTCAACGAGA 496
 DB 983 TCTTCATCTGTGTGAGAGCTTCCACCAAGAGAGACATCTGTGCACTGATCAATAAA 1042
 QY 497 GCGGCTCAGCTTTGAGATGTGTCTCTACAGCAACCTGTGATGCAAGCTGTTCATG 556
 DB 1043 CAGGACCATTAATGTGTGTATGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1102
 QY 557 CCAGGAGAGTGAATGGAGAGAGCCGAGTATTTTCCACCATCTCCAGCCGAAAATGACCA 616
 DB 1103 GCACGTCTGTGAGAGAGAGAGATTAATCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 1162
 QY 617 AGAATCTGAGAGGAGATGAGATGTGTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 676
 DB 1163 GTGACCTGAGATCTTTCAGCAATGTGTGATGTGTGATGTGTGATGTGTGATGTGTGAT 1222
 QY 677 TGATTAAGATCCCTTCGACACCTTACCATATCCCTGACTTGTATCTATCTATGTCT 736
 DB 1223 TCATCAAGATTCCTCTGACACCTTACCTGCTCTCTCTCTCTCTCTCTCTCTCTCTCT 1282
 QY 737 ATGCTTTAGCAGTGAACATCTTGTCTACTTTTGAACCTCTCAACCTGAGATGTGTCTC 796
 DB 1283 ATGCTTTGCAAGTGGGGGGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1342
 QY 797 CACCAAGGCTCA---CAACCAAGAGAGAGTGTATATCTCAAGCTGTGAGGCTTTGA 853
 DB 1343 GCATGACCATTAATGAGTGAAGACCTCTTCTATCTCTCAAAATTTGCGCTGTGGA 1402
 QY 854 AGAGAGACACACCTTCAACTCTATGTAGAGGTGCCATTTGCTGTGAGCGCAGTGGG 913
 DB 1403 AGATGACCCCAAGTGCACCTCTATGTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 1462
 QY 914 TGGAGTACCGGCTGTCTGACAGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 973
 DB 1463 TGGATATCGGCTTGTGAGGACCTTACTTTGAAAGCAGGAGAGAGCTTACGTCAGG 1522
 QY 974 CCTTGAAGTCCATTCAGATGATGACTGTCTTCAACCGTCTTCTCAAGGCGCAGAGCT 1033
 DB 1523 CCTTCAACATCAGAGAGAGACCAAAATCTCTGTTTGCATCTTTTCAAGGGGCAAAAC 1582
 QY 1034 GGAATAAATAATCCCTGATAGTCCGCTGTGATCTTCTGATCTTGAAGCAGATTAATG 1093
 DB 1583 AGTACCAACACCCCTGTATGATCTGTCTCTGTGTCTTCCCATCCGGGCGCATCAAT 1642
 QY 1094 ACCGATTAAGAGAGGCTGTGATCTTGTATACGGGCGAGAGGACAGCTGTGAGCTT 1153
 DB 1643 TGCATATCAAGAGAGGCTGTGAGTCTGTATACCAAGAGAGGACACTTGTGAGCTCACT 1702

Db 3660 GTGATACCCGCAAGGTCTGGACATCTGGAAACGCGCAGATGATTTGATTTCTCTTTA 3919
Qy 3368 ACAAGTCCAGTCCCTGCTCATCTCTCAACAAGACCAACTTCACTACTATCCCAACCGG 3427
Db 3920 ACAATGTTCAATCTTACTATCTATTAAGACCAAGTTTATCTACTAACCCCAACCGA 3979
Qy 3428 TGTTTGAGGCTTTTGTCTCTCAGAAATCTGGAGTCAAGCTTGGACGCTCATCTATCC 3487
Db 3880 CGTTTGAAGTCTCAGGCGCACTGGAAATCTTGAATCAAGAGCAGGCTTACCATCATCC 4039
Qy 3488 TAAAGGCAAGACTGATCCGCTGTGGCTGGGGGCAACGTGAAGTGAATCACTG 3547
Db 4040 TGAAGGCAAAATCTGTCTCTGTCTGTGGAGG---GGCAACTCAATCAACAG 4096
Qy 3548 TGTGTTGGAGGAGAACCGTGAACCGTGAACGTTGCAATGTCCAGGCTGTCGAGT 3607
Db 4097 TAAATGATTGAGAGACCTTGTACAGTCACTGTGTGAACACACAGCTGCTTGTGAAAC 4156
Qy 3608 CCCCCAATCTATCGCAGGCAACAAAGTATGACCCGCTGCTGGGCAATGAGTACTCC 3667
Db 4157 CTCCCAACTCAAGGCGACACAAAGTATGATTCACGTGGGCGGAAATGGTTCTTAC 4216
Qy 3668 CGGGAGTGTATCATTTGCCCCGGGACAGGCGCTGAGCTGGCGGCAATGTCAGCATCG 3727
Db 4217 CTGGCTCCGTAGGCTATCTCCGACAGGCTGTTGACCTTGCAGCATCATTCAGCATCG 4276
Qy 3728 CAGTGTGGCGGCTCTCTCATATTTCATGATGAGCGGTCTCATTTGCTTAAACGCA 3787
Db 4277 CAGTGTGGAGGCTCTCTCTTATCATGTCATCATTTGCTCTTATCCCTTAAAGCCCA 4336
Qy 3788 AGTCCCGGAAAGTACCTCAGCTGAAGGCGTGGAGATGAGATGAGCAACCTGAGT 3847
Db 4337 AGTTAGGAGAAATGACTTACACTCAAGCGGCTCAATTCGAATGAGCAACCTGAGT 4396
Qy 3848 CCCGTGGCCCTGAGTGCAGAGAACTTTTCCGAGTGCAGACGACATTCATGAGC 3907
Db 4397 CCAGGTGGGACATGAGATGCAAGAACTTTTCCGAGCTTCAAGACAGCTCAATGAGC 4456
Qy 3908 TGACCAATGACTGATGAGAGCGGATTCCTGCTGATCTATGAACTTAAACATATG 3967
Db 4457 TAACTAGATGATGATGATGATGAGAACTCTTACTGATCTGATCTGATGATGATG 4516
Qy 3968 GGGTGTCTTCCAGAGATTTGAAGACCACTCTCTCTCCGAGACCTTGAAGTCCCGGCT 4027
Db 4517 GAGTCTGTCTCCAGGATTTGAAGACCACTCTCTCTGCGGAACTGAGGGAACGGGAA 4576
Qy 4028 ACCGCAAGAGCTGTGAGAAAGGCTTGAAGCTTCCGCAAGCTCATCAACAAAG 4087
Db 4577 ATGACAGAGACAGTGTGAGAAAGGCTTGAAGCTTCCGCAAGCTTAAACAAAG 4636
Qy 4088 TGTCTCGTCTCTCTTATCCGACAGCTTGAAGTCCGAGGATCTTCCATGAGGAGC 4147
Db 4637 TGTCTGTGCTGATCTTATCCGATGAGTGAATCAAGAGCTTCTTCATGAGAGC 4686
Qy 4148 GTGGCAAGTGGCTCACTCATATGACCGTGTGAGAGCAAGCTGAGATAGCCACTG 4207
Db 4697 GTGGAAAGTGGCTCTCTCATATGACAGGCTTGAAGGCTGCTTGAATATGCACTG 4756
Qy 4208 ATGAGTGAAGAGAGCTGAGGAGCTCATTTGACAAAGAACTGGAGAGCAAGACAGC 4267
Db 4757 ATGCTCTCAAGAGAGCTCTCTGATCTCATTTAACAAGAACTTGAAGAAACAAAC 4816
Qy 4268 CTAAAGTGTCTCTCAGAGAGATGAGTCAAGTGTGAGAGAAATGCTGACCAATGTTT 4327
Db 4817 CCAAGCTGTCTCTCCAGAGATGAGTGTGAGCAGAAAGATGCTGATTAATGAGTTG 4876
Qy 4328 CTTCCTCTCTCAAGATCTCTCAAGAGATGAGTGTGAGGAGCCCTCTTCTCTCTCT 4387
Db 4877 CTTCCTCTCTCAAGATCTCTCAAGAGATGAGTGTGAGGAGCCCTCTTCAATGCTATCT 4936
Qy 4388 GTGCATCAAGAGAGATGAGAGAGGCTTGAAGCTTGAAGCTCAAGGCGAGGCGCT 4447

Db 4937 GTGCATCAAGAGAGATGAGAGAGGCTTGAAGCTTGAAGCTTGAAGCTTGAAGCT 4996
Qy 4448 ACTCTTGAAGAGAGACAAAGTCTATCCGAGAGATTTGAACAAACCTGTGCTGA 4507
Db 4997 ACTCCCTGATGAGAGACAAAGTCTATCCGAGAGATTTGAAGCTTGAAGCTTGAAGCT 5056
Qy 4508 GGTGTGAGGCGGAGACAAATGCGAAGGCGGAGGCTTGAAGCTTGAAGCTTGAAGCT 4567
Db 5057 ACTGTGTAACTCTGACATGAGAGAGAGGCTTGAAGCTTGAAGCTTGAAGCTTGAAG 5116
Qy 4568 ACACATCACTGAGTCAAGAGAGAGATTTGATGATCTTGAAGCTTGAAGCTTGAAGCT 4627
Db 5117 ACACATCACTGAGTCAAGAGAGAGATTTGATGATCTTGAAGCTTGAAGCTTGAAGCT 5176
Qy 4628 CCAAGGCGGAGAGGCTGTGAGATGATGATGATGATGATGATGATGATGATGATGAT 4687
Db 5177 CCAAGGCGGAGAGGCTGTGAGATGATGATGATGATGATGATGATGATGATGATGAT 5236
Qy 4688 TGAATTTGAAGATGAGACATCAACCAAGATTTGAAGATTTGAAGCTTGAAGCT 4747
Db 5237 TGTGTGTGAGAGAGAGACATTTACCAAAATGAGGATGATGATGATGATGATGAT 5296
Qy 4748 CACTGAGGCGGAGAGGATGATGATGATGATGATGATGATGATGATGATGATGAT 4807
Db 5297 CACTGATGATTTACAGGATGATGATGATGATGATGATGATGATGATGATGATGAT 5356
Qy 4808 CAGCTTATTAAGAGATGAACTTCACTGCTTTCAGAGCTTCAAGAGATTAATGAA 4867
Db 5357 CCTCTTAACATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 5416
Qy 4868 ACATGATCCGATTAAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 4927
Db 5417 CTTCCTTGAAGTACAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 5476
Qy 4928 CTGACTGAGAGAGTGTGATGATGATGATGATGATGATGATGATGATGATGATGAT 4987
Db 5477 CAGACTTGAAGAGGAGTCAAGGATTTGAGATCTGAGTGAAGTCAATGACATGAGGAG 5536
Qy 4988 AGAAGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 5047
Db 5537 AGAAGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 5596
Qy 5048 CCATTAAGAGGAGAGTGTGATGATGATGATGATGATGATGATGATGATGATGATGAT 5107
Db 5597 CCACAGAGGAGAGCTTGAAGAAATTTGATGATGATGATGATGATGATGATGATGAT 5656
Qy 5108 CAGACGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 5167
Db 5657 TGCACCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 5716
Qy 5168 AGGCTGATTAAGATGAGGATTTGATGATGATGATGATGATGATGATGATGATGAT 5227
Db 5717 AGGAGAGAGAGAGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 5776
Qy 5228 TGCCTGAGGATTTGAGTCAATGATCAAGAACTTGTGATGATGATGATGATGAT 5287
Db 5777 TGCACCTTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 5836
Qy 5288 AGAAGAGATCAAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 5347
Db 5837 AGGAGAGATCAAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 5896
Qy 5348 CCAAGTCAAGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 5407
Db 5897 CCAAGTCAAGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 5956
Qy 5408 ACATGCGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 5467
Db 5957 ATATCCAGATTAAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 6016
Qy 5468 CCATGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 5527
Db 6017 CCATTAAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 6076

Db	1177	GAACATTAAGAGCGCATCCAGTCGTCTGCTACCGTGTGTAAGGGCAAGCTCTCCCTGCAGT	1236
Oy	1155	GCTTAAGATGAAGACATCCCTTCGACAGAGGCGCTCTTAACCATTAACGATAACTCTG	1214
Db	1237	GCTCTCAACAGGAGCTGGGCTGTGATCACTCGCCCTCGAGATCGATGACACTTCTG	1296
Oy	1215	TGCGCTGGACATGAACTGCTCCCTTGGAGTGTCCGACATGAGTGGGAAATCCCGTCTT	1274
Db	1297	CGGGCAGGACTTCAACGCGCCCTTGGGGGGCAGACTGCATTAAGAGGGAGCGCCCTGTT	1356
Oy	1275	CACGGAGACAGGAGCCCGCATGAGCGTCTGTCATGCGATATGTATCAAGAACCACTCTCT	1334
Db	1357	CTGGACAGGATGTATGGCTTGACCGCGCTGGCTTCTATATACTATCTGGGGCCGCACTGT	1416
Oy	1335	GGCCTTTGTGGGACCAAAAGTGGCAAGCTTGAAGAAAGATCCGGGTGATGAGCCCA---	1390
Db	1417	GGTATTCGCCGGCAGCGCAAGTGGCGCGATCCGCAAGATCCTGGTGGACCTCTCAAAACC	1476
Oy	1391	-----GGGGCAACGCCCTCCAGTATGAACGGTGAAGGTGGTGAAGCCCGGCCAGTCT	1445
Db	1477	CGGTGGCGGCGCTGCGCTTGGCTACGAAAGCGTCGTGGCCCAAGAGGGCAGCCCATCT	1536
Oy	1446	CCGGAAATATGCGCTTCTTCAAGAGCACAGACCAACTTACATATGTCAGAGAGGAGCT	1505
Db	1537	GCGAAGCTCGTCCCTCAGCCCCAACCAACAGTAACTCTTAACGCATATGACCGAAGCAGGT	1596
Oy	1506	CACCAAGATCCCTGTGGAGTCTGTGGTCAGATATAGAGCGTGGCGGAGAGTCCCTGGCTC	1565
Db	1587	GACGGGGTGCCTGTGGAGAGCTGTGTCAATACACGTCTGTGAGCTGTGTCTGGGGTC	1656
Oy	1566	AGGCGACCCCACTGTGGCTGT	1625
Db	1657	ACGGGACCCCACTGTGGCTGT	1716
Oy	1636	TGAGGGGTCCAAAGAGCGCCCGAGTTTGCCCTCGAGATTAAGAGAGTGTGTCCGGCTAC	1685
Db	1717	TGAGCGACAGACGAGCGCCCAAGCGCTTTGTCTGGGACCTGCTGCAGTGTGTGCACTGAC	1776
Oy	1686	GGTCCATCCCAAAATATCTCCGTCTCAAGTCAACGTGC---TGCTGTCTTGGAGAC	1742
Db	1777	TGTGAGCGCCCGCAATGT	1838
Oy	1743	GTAACAATGCCGAGCTGTACAGCTGGGGGTCAACTGACCTTTAGACCTGTCAAGAT	1802
Db	1837	CTGGAACGTGCTTACCTCTCAGCTGGGCTCACTGCTCTCTTCAAGAACTTCACGGAATC	1899
Oy	1803	GGATGGGTGTGTGTGGGCAATCAGATCCAGTCTACTTCCCTGCAGCCAAAGAGTGC	1862
Db	1897	TGAGAGCGTCTCGAGAGATGAGCGGATCCACTGCGCTCAACCTCCCGCCGGAGGTGC	1956
Oy	1863	CCGAGTCA-----TCACAGAAATGGGGACCAACATGTGTAACGTTCAGCTCAATTC	1914
Db	1957	GCCCATCAGCGGGGGCCAGGGTGAAGGAGACCAAGGGGTGTGAATCTCTACTTAATGC	2016
Oy	1917	AAAGAGACCCGGCATGACCTTGTCCAGACACAGCTTTGTCTTCAAAATGCAAGCTCCA	1978
Db	2017	CAGGAGACAGGGAAAGATTGTGCTGTGAGACTTGGCTTCAACTGCACTGCAAGGTCCA	2076
Oy	1977	CAATTG---TGCGTGTCTGTGTGAGAGTCCATAACGGTGCACACTGGTGAATTAACG	2033
Db	2077	CCAAGTGAAGCTGCTGTCTGTGTCAACGGCTCTTTCCCTGCCACTGTTGAATAACG	2136
Oy	2034	GCATGTCTGACCCCAATGACCCCAAGACCTGTCTCTTCAAGAGGCGCAATGAAGCTGCC	2099
Db	2137	CCACGTGTGACACACAACAAGTGTGACCTGACCTTCTGAGAGGGCCGTGTAAAGTGTTC	2196
Oy	2094	CGAGACATGCCCCCAAGCTGTCTGAGTGGACAAGATCCGTGTGCCCGCTGAGAGTATCA	2153
Db	2197	TGAGGACTGGCCCAAGATCTGTCCCTCCACGCAAGATCTACGTGTCCAGTGGAGTGTAA	2255
Oy	2154	GGCTTACAGCTGAAGGCCAAGAACTCCCCAGGCCCAAGTGTGGAGAGCGGTGTAAGA	2213
Db	2257	ACCCATCACCTTGGCGCAGCAAGTGTGCAACAGCAAGTGAAGGCAAGCGTGAATATGA	2316

QY	2214	ATGATCTCTCAAACTTTCAGGGACGCGACGACGAGTGTCCGCGCTGTGACTTCAACACTC	2273
Db	2317	GTGCTCTTTCACATCCCGGAGACCGCGCCGTGTCAACCGCTTGCTTCAACGACTC	2376
QY	2274	CAGCGTACAGTGCAGAACTCTTTATTTCTTATGAAGGATGGAATGCACACTCTCC	2333
Db	2377	CAGCTTGCACTGCGAGAACTCTCTGTAATCTCTTACGAGGGAAAGATGTACGACTCTGCC	2436
QY	2334	CGTGAAGTTGACAGTGTGTGGAATGGGCACTTCAACATTGACCAACCACCTCGAATAA	2393
Db	2437	AGTGAACCTTGCACTGTGTGTGGAAACGGCAACTTTGTCAATTGACCAACCAAGAACATCCA	2496
QY	2394	AGTTCACTTCTCAACAGTGTGAAGCCATGTGAGAGCTGTGGGTGTGCTCAAGGCTGA	2453
Db	2497	GGGCACTCTTAAACAAGTCCCGGCGCTGTGCGCGAGAGCTGTGGGCTCTGTCTCAAGGCGGA	2556
QY	2454	CCCAAGCTTGCAATGTGAGCTGGTGCACAGGGCCCAAGGCGACAGTGCACCTGTGCGCCGACTG	2513
Db	2557	CCGCGCTTGAAGTGTGGAATGTGTGTGTGTGGCGCGAGCGCGCTGTCTCTCTCGACACCACTG	2616
QY	2514	CCCTGCCCA--GGAGAAGCAAGTGTGAGCTGTGTGTGTGTGCGCAAAAGCAAGTGCACAA	2570
Db	2617	CGTGTCCGACACACTGTGATGTGGAAGCAAGCGCGTCAAGGCGAGCACTGTCTGCACCGGA	2676
QY	2571	CCCCCGATCAAGAGATTAATCCCGGTGACAGGCGCCCGGGAAGGGGGACACAAAGTCAAC	2630
Db	2677	CCCCAAGATCTTCAAGCTGTGTCCCGGAAGGCGCCCAAGGCGAGGCGGACACGGGCTCAC	2736
QY	2631	TATCCAGGGGAGAACTCTGGGCGCTGGAATTTCCGCAATCGCTCCCATGTCAAAGTTTC	2690
Db	2737	TATCAAGCGAAGAACTTGGGCGTGGCAATTCGAAGAAGTGTGTGGGCGTGGCGCGTGGG	2796
QY	2691	TGGGTGTGAGTGTGACGCCCTTTAGTGAATGTGTTCATCCCTGCAABAACAGATCGTGTGA	2756
Db	2797	CAAGTCTGTGTGACGCCCTTGTGAGAGCGAGTATCATAGTGTGGAGCAATGTCTGTGA	2856
QY	2751	GATGGGAGAGCCCAAGCCC--AGCCAGCAATGACAGCTTTCGTGGAATCTGTGCTGTGCTGT	2807
Db	2857	GATGGGGAAGCCACGCTCCGTGTGCGTGCATGAAGCCCTGGTGAAGGTGTGTGTGCGGGA	2914
QY	2808	GTGTCTGAGTGAATTCATGAGCCGCTCTCTCAAGCTCTTATTACTTCATGAACAGCTCT	2865
Db	2917	CTGTGTACACACTACCGCGGCTGTGTACCCCAAGCGCTTCACTTCTGTGAACCAACTT	2978
QY	2868	CTCAGATCTGMAAGCCAGCGGGGCCCATGTCCGAGGGAAGCCCAAGTGCACATCAAGG	2927
Db	2977	CTACCGTGTAGCCCTCCGTGTGGCTGTGTCAAGGGGGCACTGGATGTGCATCAAGGG	3038
QY	2928	CACCAACTGTAATCCCGGAAGCAACGTGTGTGATTTTGGAAAAGCAAGCCCTGTCTT	2988
Db	3037	AAGCACTGTGAACGAGGCACTGATGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTCTT	3099
QY	2988	CCAAGCGCATCTCTCATCTTCACTTGTCTGTGACACCAACATCTCTCAGATGAGTGTGA	3047
Db	3097	CTCTGTGTCCAGAGGAACCTCCGTGAATCCGCTGTGTGACACCCCGGGCAGAGGCC	3156
QY	3048	GATGAAGTGTGGTGCAGGTGTGACAGGGC-----AAGATCCACCGAGACTGTGT	3099
Db	3157	TGGCAGGCTTCCATCATCATCAACATCAACCGGCCCAAGTCAACCAACTTGAAGTGA	3218
QY	3099	CTTTCAGTATGTGAAGAAGCCCAACATGTGTGGAATGAGCAATGAGCAATTTGAG	3158
Db	3217	GTACACCTACACGAGAGACCCCACTCTGTAGATGCAACCCGAGTGAATCAACAG	3278
QY	3159	TGGAACACACCCATCGCTGTATGAGGGGACCACTGTGAACCTCAATACGAACCCCAAT	3218
Db	3277	CGGTGGGACCTCTGTAGGTCAACAGGACCAACCTGTGCACCTGTCCGTGAACCCCAAT	3333
QY	3219	CCGTGTCCAGATGAGGGAAGAGCAACATCAATCTGTGAGGTTTGAAAGCTTACTGA	3278
Db	3337	CGGGGCAAGTATGGAAGCAATTGAG---AGGAGAACTGTGTGTATCAATGACACAC	3393

QY 3279 GATGACTGTCAAGGCGCCCGCTGTGCTGTGAGTCCCTGACCAACCAAGTCAAGCTTGACCGA 3338
DB 3394 CATGATATGCCGCGCCCGCTGTGCTGTGAGTCCCTGACCAACCAAGTCAAGCTTGACCGA 3453
QY 3339 GAGGCGCCGAGAGATTGGCTTATCCTTGACCAACCGCAAGTCCCTGCTCATCTTCAACAA 3398
DB 3454 GCGGCGGAGATGAGCTGAGCTTGTCTGATGACCAAGTGGCTCTCCCTGCTTGTGCTCACTC 3513
QY 3399 GACCAACTTCACTTATCCCAACCGGATGTTGAGGCTTTGGTCCCTCAAGAACTCT 3458
DB 3514 CACCTCTTCTCTCTACTACCTGACCCCGTACTGAGGCACACTCAACCCCACTGAGCTCTCT 3573
QY 3459 GAGGCTCAAGCTTGACCGCCATCATCTTAAAGGCAAGAACTGATCCCGCTGTGAGC 3518
DB 3574 GAGGCTGAAGCCGAGCTCCCACTCATCTTCAAGGCGGAACTCTTGGCACT--GC 3630
QY 3519 TGAGGGCAAGTGAAGTGAAGTCACTGCTGTGTTGGGAGAGCCGTGACCCGTGAC 3578
DB 3631 ACCCGCAACTCCCGACTCACTACACGCTGCTCATCGGCTCCACACCCCTGTAACCTCAC 3690
QY 3579 CGTGTCAATGTTCAGCTGTCTGCGAGTCCCACTCATCGGCAAGCAAAAGTAT 3638
DB 3691 CGGTGTGAGAGCGCACTGCTGTGCGAGGCGCCCAACTCACTGAGCAGCAAAAGTATC 3750
QY 3639 GGCCTGTCTCGGTGATGAGTACTCCCGGAGATGTTGATCACTTGGCCCGGACAGCC 3698
DB 3751 GGTGCTGCAAGTGTGCTTGAATTTCTGCGCAGGACACTGCAAGTGTATCTCGGACAGCT 3810
QY 3699 GCTCAAGCTTCCCGCATGTGCAAGTGCAGTGTGCGGCTCTCTCATATTTTCAT 3758
DB 3811 GCTGACGCTGCTGCTATGTGGGCAATTGGCGAGCGGAGCTCTCTGCTCTGCTAT 3870
QY 3759 CGTGGCGGTTCATTTGCTTAAAGCAAGTCCCGCAAGTGAAGTCAAGCTGAGAGG 3818
DB 3871 CGTGTGTGTCTATGCTTCAAGCGCAAGTACAGAGTGTGACCGACACTCAAGCG 3930
QY 3819 GCTGCAATGCAAGTGAACAACCTGAGTCCGTTGTGCTTGAAGTGAAGAAAGCTT 3878
DB 3931 GCTGCAAGTCCAGATGACCACTGAGTCCCGGCTGCGCTTCAATGCAAGAAAGCTT 3990
QY 3879 TCCGAGCTGCGACAGCAATCATAGTGTGACCAATGACCTGATGAGCCGAGATTC 3938
DB 3991 TCGAGAGCTGCAACAGCACTCCAGAGCTGACCAATGACCTGAGCGGTGCGGCACTCC 4050
QY 3939 GTTCCGAGCTATAGAACTTACACCACTGCGGCTGTGTTCCAGAGATTTGAACCAACC 3998
DB 4051 CTTCCTTGACTACCGGACATATGCAATGCGGAGTCTTTCCTGGGATCGAGAACACACC 4110
QY 3999 TGTCTCCCGGACCTTGAAGTCCCGGCTTACCGGCAAGAGCTGTGAGAAAGGCTGAA 4058
DB 4111 TGTGCTCAAGAGATGAGAGTTACAGGCA-----TGTGAGAAATCTCTGAC 4158
QY 4059 GCTCTTCGCGCAGCTCATCAACAAGAGTCTCTGCTCTCTTATCCGACCGCTTGA 4118
DB 4159 ACTGTTCCGCGCACTGCTGACCAAGAAAGCACTTCTCTGCACTTCAATCGGACGTTGA 4218
QY 4119 GTCCCAAGCTGATTTCTCATGCGGACCGTGGCAAGTGGCTTCACTCATGACCT 4178
DB 4219 GGGACAGCGCAGCTTCTCATGCGGACCGCGGAAATGTGGCTTCCCTCATCATGACGCGC 4278
QY 4179 GCTGCAAGAGCTGAGTGAAGCACTGATGTGCTGAGAGCACTGACCGCACTCAT 4238
DB 4279 CTTGCAAGGCGAGATGGAATACCGCAAGGCGTGTCTCAAGCACTCTTCCGACTCAT 4338
QY 4239 TGAAGAAGCTTGAAGAGCAAGAACCACTTAAGCTG---CTGCTCAAGAGAGCTGAGTC 4295
DB 4339 CGAGAGAACTTGAAGAGCAAGAACCACTTAAGCTG---CTGCTCAAGAGAGCTGAGTC 4398
QY 4296 AGTGGCTGAAGATGTGACCAATGTTGTTACTTTCTCTCTCAAGTCTCTCAAGGA 4355
DB 4399 GGTGGCAAGAGATGTGATCTGATGCTGATCACTTCTCTTGTATAGTTTCTCAAGGA 4458
QY 4356 GTGTGTGGAGAGCCCTCTTCTCCCTGTTCTGTGCAATCAAGCAGAGTGAAGAGG 4415

DB 4459 GTGGCTGGAGAGCCGCTGTTCATGCTGATCGCATCAAGCAAGTGAAGAGG 4518
QY 4416 CCCCATTTAGCGCATCAAGGCGAGGCGGCTACTCTCTTGAAGGAGACAAAGCTATCCG 4475
DB 4519 CCCCATTTAGCGCATCAAGGCGAGGCGGCTACTCTCTTGAAGGAGACAAAGCTATCCG 4578
QY 4476 CCAAGATTTGATCAAAACCTTGTCTGAGCTGTGTGAGGCGGAGCAATGCCAAGC 4535
DB 4579 GCAAGCAATTGACTCAAGCACTGACCTGAACTGTGTGAACCTTGAGATGAAATGC 4638
QY 4536 CCCCAGGCTCCAGTAAAGATCTCACTGATGACCATCACTCAAGTCAAGAGAGAT 4595
DB 4639 ACTGAGTGCCTGAGAGGAGGCTGAGCTGTGACAGGTCAACCGGCAAGGAGAGCT 4698
QY 4596 TCTGATGCCATCTTCAAGATGTGCTTGTCTCCACCGGCGCAAGCTGCAATGGA 4655
DB 4699 GCTGGAAGCTTCTTCAAGGAGGCTGCTTACTCCAGCGGCGCAAGGCGCGGACATGGA 4758
QY 4656 TCTGAGTGGGCACAAGAGTGGGCGAGAGTGAATCTTCAAGATGAACATCAAC 4715
DB 4759 CTTGAGTGGCGCGAGGCGGCGATGCGGCGATCTCTGCAAGAGAGAGAGCTCAAC 4818
QY 4716 CAAGATTTGAAGTATTTGAAGCGACTGAAACACTGGCCCACTTCAAGGTGCCAGATGG 4775
DB 4819 CAAGATTTGAAGATTTGAAGAGGCTGAAACAACCTGGCTCACTACCAAGGTGACAGCG 4878
QY 4776 TTCCGTGGGCAATTAGTTCGAAGCAGTGAAGAGCTTATTAAGCAAGTGAACACTCAC 4835
DB 4879 GTCTTGTGTGCACTGTGCTCCAGAGAGCTGCTGCTTCAACATCTCCAACTCTCCAC 4938
QY 4836 CGTCTCCAGACCTCAAGCAATTAATTAAGAAATGATCTCGGTAACGCGGACGCCCGA 4895
DB 4939 CTTCAACAG---TCCCTGACAGATGAGAGAGATCTCTGCAAGCGGCGAGCGCCGCA 4995
QY 4896 CAGCTTCCGCTCAAGCACTTATGATCACTCTGACCTGAGAGTGAAGTCAAGATGTG 4955
DB 4996 CAGCTTCCGCTGCGGACGCGCCATGATCAAGCTGAGCTGAGAGCGGCACTCAAGCTGTG 5055
QY 4956 GCACTTGAAGAACCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAG 5015
DB 5056 GCACTTGAAGAACCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAG 5115
QY 5016 GGTGTCTGAAATTTACTGACCGCACTCTCTGCGCACTA---AGGCACTTGAAGAGTT 5072
DB 5116 GGTCTGGAAGATCTTACTGACCGCTTACCTGCGCACCAAGCAAGGCACTTGAAGAGTT 5175
QY 5073 TGTGATGAACCTCTTGAAGCACTTCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAG 5132
DB 5176 TGTGAGGACCTGTTTGAAGCACTTCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAG 5235
QY 5133 GGCATCAAGTCAATGTTTGACTTCTGATGAGCAAGCTGATTAATGCAATTCATGA 5192
DB 5236 GGCATCAAGTCAATGTTTGACTTCTGATGAGCAAGCTGATTAATGCAATTCATGA 5295
QY 5193 CCGGCAAGTCCGCAATCACTGGAAGAGCAAT---GCTGCGCTGAGAGTTTGGGCA 5249
DB 5296 TGTGAGCTGCGCACTTGAAGAGCACTGAGAGCACTGAGAGCTGCGCTTCTGAGTGA 5355
QY 5250 CATGATCAAGAACCGGAGTTTGTGATCATCATGAAGAACATCAAGAGCTG 5309
DB 5366 CGTATCAAGAACCGAGTTTGTGATCATCATGAAGAACATCAAGAGCTG 5415
QY 5310 CCTCTGTGTGTGTGCTCAAGCTTCAAGAGCTGCTGCTGCAAGCTCAAGAGCAAGCTGAG 5369
DB 5416 CTTGTGCTGTGTGTGCTCAAGCTTCAAGAGCTGCTGCTGCAAGCTCAAGAGCTGAG 5475
QY 5370 CAAGAGCTGCGCTTCAAGAGCTGCTGATGCGAAGAGCAATCCCAAGCTTCAAGAGTTG 5429
DB 5476 CAAGAGCTCAAGCTTCAAGAGCTGCTGATGCGAAGAGCAATCCCAAGCTTCAAGAGCTG 5535
QY 5430 GGTGGA---GAGATTAATCAAGCATGAGAGAAATGCCAGGCAATCAAGAGCAAGCAT 5486

Db 873 CCTCCTTGGTGTACTTCTGACCGCTGACGTGAGACCCAGAGACGCTGTTGACACAG 932
Qy 812 CCAAGAGACGAGTGTATACATCCAGCTGTGTAGGCTTTGCAAGAGACACAGCCTTCA 871
Db 933 CGGCGCAAGAAATTTTTCAGCTCCAGATGCTGTGCGATGTGCGCGGAGACTAGAGTTCT 992
Qy 872 ACTCCATATGAGAGTGGCCATTTGGCTGTAGGCGAGTGGGGGTAGTACCGCTGCTGCG 931
Db 993 ACTCAATAGTGAATTCCTCATGAGCTGCTCTGTGGCGCGGCTGAGTACCGCTTGTGTC 1052
Qy 932 AGGCTGCTTACTGTCCAAAGCGGGGCGGTGTTGGCAGAGCCCTTGAATTCATCCAG 991
Db 1053 AGAGCGCCCACTGGGCCAAGCCTGGCTGTGCTGAGCCAGAGCCTTGGGCGTGCAGCTG 1112
Qy 992 ATGATACCTGCTCTTCAACCTCTTCTCCAGAGGCGCAAGAGGAAATGAAATCCCTGG 1051
Db 1113 ATGAGAGAGCTCTTCAACCATTTCTCTAGGCGCGAAGAAACGGGCGACCCACCC 1172
Qy 1052 ATGAGTCGGCCCTGTGATCTTTCATCTTGAAGCAGATAATGACCGCATTAAGAGCGGC 1111
Db 1173 GGCAGACCATCTCTGTGCTTTCAACCTCAGCAACATCAATGCGCCATCGGCGCGCA 1232
Qy 1112 TGCAGTTTGTACCGGGCGAGGGGACGCTGGAACCTGGCTGTGCTCAAGGTGAAGACA 1171
Db 1233 TCCAGTCTGTCTTCTGTGGGAGGACCTGTGCTGCTGCTGTGCTGAACAAGAGAC 1292
Qy 1172 TCCCTGACAGCAGTGGCTCTTAAACCATTAACGATTAATTGTGGCTGGAATGAATG 1231
Db 1293 TGGCTGTGATCAACACCCCATGAGATCAACGGCAACTTCTGTGGCGGTGTGTTGAAC 1352
Qy 1232 CTTCCTCTGGAGTGTCCGACATGTGTGCTGGAATTCCTGCTTTACGAGAGACAGGACC 1291
Db 1353 AGCCTGTGGAGGCGCTGATGTGATGAGGGGCTGCTGCTGTGCGCAACAGACCGAG 1412
Qy 1292 GCATGACGTCTCATGCAATATGTATACAGAAACCACTCTGTGGCTTTGAGGACCA 1351
Db 1413 GCATGGCCAGGTGTGGCGCCTTACACTTACCGCAGACCTGTGTGTCTTACATTGGCAGC 1472
Qy 1352 AAAGTGCAGAGCTGAAGAAAGATCCGGGTGATGAGACCCAGGGGCAACGCTTCAGTATG 1411
Db 1473 GCAGCGGAGCTTGAAGAAAGTGTGGGTGATGTGCTTCAGG---ATGCGCACTGTATG 1529
Qy 1412 AGACGGGACAGGTGTGAGACCCCGGCGCACTCTCCGGGATATGAGCTTCCAAAGAAC 1471
Db 1530 AGACAGTCCCGGTGTGTAGTGTGAGACCCCATCTCCGAAGCTGTCTTCAAGCCGAGCC 1589
Qy 1472 ACAGGCAACTCTATCATATGTCAAGAGGAGCTCAACAGATCCCTGTGAGTCTGTG 1531
Db 1590 ACCGGACATCTATCTCTGTAGTGAAGAGAGTGAAGCCAGCTCCGGGTGAGAACCTGTG 1649
Qy 1532 GTCAATATCAGACTGTGGGAGATGTCTTGTCTCAAGCGAACCCCACTGTGTGCTGTG 1591
Db 1650 AGCAGTACAGAGTGTGAGCCTGTGCTGTGGCTCCGAGACCGGCACTGTGTGTGTG 1709
Qy 1592 TGGTGCACACAGTGTGACCCGAGAGAGCGGTGTGAGCGGTCAAGAGACCCCGAGGT 1651
Db 1710 TGTGTGAGACAGAGTGTGTGCGAGAGGGGCTGTGTGGGCGCTGTGCCCCACAGGCT 1769
Qy 1652 TTGGCTGTGAGTGAAGAGATGTGTGCGGTACAGCTCATCTCCACATATCTCGTCT 1711
Db 1770 TTGCTGAGGAGCTGAGGAAATGTGTCCAGGTGTGGGTTCGGCCCAACATGTGTCA 1829
Qy 1712 CTCAATACACAGTGTGTGTCTGTGAGAGCTACATGTCCGAGCTGTGAGCTGCG 1771
Db 1830 CTGTACCTGTGGGTGTGAGCTGTACCGTCAACCTGTCAACGTGTCCAGACCTCAAGTGGGG 1889
Qy 1772 TCAACTGTCACTTTGAAGACCTGTCAAGATGTGATGG---CTGTGTGTGGGCAATCAGA 1828
Db 1890 TGAAGTGTGGCTTGTGAGGCGGCGGAGAACGAGGCGGTCTGTGCGCTCCGGTGAAC 1949
Qy 1829 TCCAGTCTACTCCCTGTGAGCCAAAGAGAGTGTGCCCCGATCATCAGAGAAATGGGAGCC 1888

Db 1950 TGTCTGCCCCCTCAACCTCCCTCCAGGAGCTCCAGACTTTACAGAGGGGAGTGTGGGCA 2009
Qy 1889 ACATATGTGTACAGCTTTCAGCTCAAAATCAAGAGACCCGCAATGACCTTCCAGACCA 1948
Db 2010 CCGGCACTGTGCGGTGTGACACTTCTCTCAAGAGAGAGCGGTGAGGTGTGGGTGTG 2069
Qy 1949 GCTTGTCTTCTCAATATGACGCTTCCAAATTCAGTTCGTGTCTGTGTGAGAGTCAAT 2008
Db 2070 ACTTGTCTTCAATATGACGCTCTTCCAGTGTGATGTCTGTGTGTGAGACCTT 2129
Qy 2009 ACCGTGTCACTGTGTGAATATCCGAGATGTGTGACATGACCCCAAGACCTGTCTCT 2068
Db 2130 ACCCTGTCACTGTGTGAATATCCGAGATGTGTGACATGACCCCAAGAGTGTCTCT 2189
Qy 2069 TCCAGAAAGCGAGTGAAGCTGTGCGGAGACTGCCCCAGCTGTGTGAGTGAACAGA 2128
Db 2190 TCCAGAGGCGAGGTCCACAGCCTGTAGGGGTGTGCTGTGATCTGTCCAGTGTGGGACC 2249
Qy 2129 TCTGTGTGCGGTGTGAGTGTATCAAGCTTATCAAGCTGAAGGCCAAGAACTTCCCTAGC 2188
Db 2250 TCTGTATCCCGTGTGGGTGTATGACGCTCTTACTGTGCGGCTAAGAACTTACTAGC 2309
Qy 2189 CCGAGTGTGGGACGCTGTGCTACGAATGCAATCTTCAAGGCGAGGAGAGAGAG 2248
Db 2310 CGCAGTGTGGGACAGAAACTATGATGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 2369
Qy 2249 TGGCGGCTGTGCTTCAACAGCTCCAGCGTACAGTGTGTGTGTGTGTGTGTGTGTGT 2308
Db 2370 TGGCTGT 2429
Qy 2309 AAGGATGTGAGTCAACAACTGTGAGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 2368
Db 2430 AAGGT 2489
Qy 2369 ACATTTGACAAACCACTCAGATTAAGTTTCACTTCAAGTGTGTGTGTGTGTGTGTGTGT 2428
Db 2490 CCAATGACAAAGCTCCAGGCTCCAGGCTCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 2549
Qy 2429 GCTGT 2488
Db 2550 GCTGT 2609
Qy 2489 GCGAGT 2548
Db 2610 ACAAGT 2669
Qy 2549 GTGCAAAAGCAATGTCAAAACCCCGCATCAAGAGATTAATCCGCTGTGACAGCCGCC 2608
Db 2670 AGAAGGACACCGGTGTGAGCCACCTCGCATCAAGAGATCAACCTCTGTGTGTGTGTGT 2729
Qy 2609 GGAAGGGGGGCAACAGTCAATCCAGGGGAGAACTGTGGCTGTGAATTTGTGCAGCA 2668
Db 2730 AGAAGAGAGGACCCGGGTGTGACATGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 2789
Qy 2669 TCGCTTCCATGTAAAGT 2728
Db 2790 TGGGCTGT 2843
Qy 2729 CTGCAAG----- 2735
Db 2844 GTGCTGTGAGGT 2993
Qy 2736 -----ACAATGT 2776
Db 2904 GACGCTCTGTGAGCCTTGT 2963
Qy 2777 ATGAGGCTTGT 2836
Db 2964 CGCGGGGCGCGT 3023
Qy 2837 CACAGCTTATTAATTGT 2896
Db 3024 AGCAGGT 3083

QY 2897 TGTCCGAGAGGACCAAGTACATACAGGACCAACCTGAAATCCGGAAGCAACGTGG 2956
Db 3084 CGTCCGAGGGGACACAGGGCTTACATCTCAGGAGCTCTGGATCTCTGGACAGAGGGTCA 3143
QY 2957 TGGTATGTTTGAAGAAGCAAGCCCTGTCTCTTCCACAGGCGATCTCCATCTCAATGTCT 3016
Db 3144 CAGTACTGTGAGGACAGCGAGTGCAGTTTGTAAAGAGATGCCAAAGGCGATCGTGT 3203
QY 3017 GCACACACCATCTCTCAGTATGAGTGTAG--AGATGAAGTGTGCTGACAGGTGACA 3073
Db 3204 GCAATCTCACTCTCTCCACCCCTGGGCCACAGGACAGCCCATCACTTGTGCATGACC 3263
QY 3074 GGGCCAGATCCA--CAGGACTGTGTCTTCAATATGTGAGAGACCCCAACATCGTGC 3130
Db 3264 GGGCTAACATCTCCAGCCCCGGGCTCATCTACCTACATCAAGAACCCACCGTCAACC 3323
QY 3131 GGATTGAGCCAGATGAGACATTTGTCAGTGAAAACACCCATGCGTATGGGGAGACC 3190
Db 3324 GCGTTGAGCCCACTGAGACATCATATGAGAAAGCTGCACTCATCTGTGAGTGGAGACC 3383
QY 3191 ACTTGACCTCATACAGAAACCCCAAGTCCGTGCAAGCATGAGAGGAGAGACATCA 3250
Db 3384 ACTTGCTAGGTTCCAGAGAGCCCGGGTCCGTGCAAGTACCGGCGCATGAGACACCA 3443
QY 3251 AATCTGTAGAGTTTGAACCGCTACATGAGATGACCTGTCAAGCGCGCCGCTCGTCTGG 3310
Db 3444 ATACATGCCAAGTACAGACACTGCACTGTGTGTAAGGCCCGGCACTTTTCTTG 3503
QY 3311 GTCTGACCAACAGTACAGACTGACCGAGAGGCCGAGAGATTGGCTTCACTCTGACA 3370
Db 3504 GGGGCCCCAGCCTGGGCGCAAGGAGACACCTGATGATGTTGGCTTCTGCTGAGACC 3563
QY 3371 AGGTCCAGTCCGTCTCATCTCTCAACAAACCAATTCACTCACTCACTCACTCACT 3430
Db 3564 AGGTGAAACGGCCCGCTCTCAACCGCTCTCTTCACTCACTCACTCACTCACTCACT 3623
QY 3491 AAGGCAAGAACTGATCCCGCTGTGTGCTGGGGGCAAGTAACTGAACTCACTGCTGC 3550
Db 3684 AAGGCAAGAACTGAT--TCCCGGGCAGCGGAGCTCCGCTCACTCACTCACTGCTGC 3740
QY 3551 TGGTGGGAGAGAGCGGTGCAACGTGACAGTGTCAAGTGTCCAGCTGCTGCGAGTCCC 3610
Db 3741 TGAATGAGAGCCAGCCGTGTGTGCTCACTGTCTCGACACAACTCTGTGCGACTAC 3800
QY 3611 CCACCTCATCGGACAGGACAAAGTATGAGCCGTGTGCTGAGTGTGAGTGTGCTGAG 3670
Db 3801 CCAAGCAGACTGAGCCGAGCCGTGTATGATGCTGATGAGTGTGAGTGTGAGTGTGAG 3860
QY 3671 GAGTGTGTATCATGTGCCCCGAGACAGCCGCTGAGCTGCCCCGCACTGATCGAG 3730
Db 3861 GCACTCTGACATCTGAGCAGCGGAGCGGCTCACTTACCGGCACTGATGGGGCTGGGG 3920
QY 3731 TGGCTGCGGCTCTCATCATTTTCATGTGAGCGGTGCTCATGCTTAATTAACGCAAGT 3790
Db 3921 CCGGAGGAGTGGCTCTGTGTGTGCTGACATCAACCGTGTGTGCTGAGTGTGAGTGTGAG 3980
QY 3791 CCGGCAAGTATGATCTCACTGAGAGCGGCTGACATGAGTGAAGTGAACCTGAGAGTCCC 3850
Db 3981 CTGAGAGCCGAGCCGTACCTTCAGCGCTGTGAGTGTGAGTGTGAGTGTGAGTGTGAG 4040
QY 3851 GTGTGGCCCTGAGTGAAGAGAGCTTTGCGAGCTGACAGCGACATTCATGAGTGA 3910
Db 4041 GTGTGGCCCTGAGTGAAGAGCTTTGCGAGCTGACAGCGACATTCATGAGTGA 4100
QY 3911 CCACTGATCTGATGAGAGCGGAGTTCGCTTCTGATGATTAAGTAACTTAACCAAGCGGG 3970
Db 4101 CTAAACCAATGAGCGAGTGTGAGATCCCTCTCTGAGTCAACGGACTTAACGCTGCGCG 4160

QY 3971 TGTCTTCCAGGAATTGAAGACCACTCTGTCTCCGAGACCTTGAAGTCCGGGCTTACC 4030
Db 4161 TGTCTTCCCGGACATCGAGGCCCAACCCGCTCTCAAGAGAGCTGATACGCCCAAC- 4219
QY 4031 GCGAGAGAGGTGTGAGAAAGGCTTGAAGCTCTTCCGAGCTCATCAACAAAGGTGT 4090
Db 4220 -----GTGAGAAAGGCCCTGCGCTCTTGGGAGAGTGTGCAACACCGCGCT 4268
QY 4091 TCTGTGTCTCTTCACTGAGCAGCTTGAATGCCAGGTAGTCTTTCATGCGGACCGTG 4150
Db 4269 TGTGTCTTACCTTCACTCAACGCTGAGAGCCAGAGCGCTTCTCATTGCGGACCGCG 4328
QY 4151 GCAACGTGGCTCACTCATATGACCGGTGCTGAGAGCAAGCTGAGTACGCGCATGTAG 4210
Db 4329 GACCGTGGCTGTGCTCAACGTGTGAGCCCTGTGAGAGCCGAGCTCACTATGCGAGGCG 4388
QY 4211 TGTGAAAGAGCTGTGAGCCGACCTTCAATGACAAAGACTGAGAGCAAGAACCACTTA 4270
Db 4389 TGTCAAGCAAGTGTGAGCCGACCTTCACTGAGAAAGACTGAGAGCAAGAACCACTTA 4448
QY 4271 AGCTGTGTCTCAGAGAGACTGAGTCAAGTGTGAGAGAGTGTGAGAGTGTGAGTGT 4330
Db 4449 AGCTGTGTCTCAGAGAGACTGAGTCAAGTGTGAGAGAGTGTGAGAGTGTGAGTGT 4508
QY 4331 TCTCTCTTCAAGATTCCTCAAGAGTGTGAGAGAGCCGCTTCTCTCTGTTCTGTG 4390
Db 4509 TCTCTCTGATTAAGTGTGAGAGAGTGTGAGAGAGCTCTCTCTCTGTTCTGTG 4568
QY 4391 CCACTAAGCAGAGATGAGAGAGGCGCCCAATTAAGCTCAACGCGGAGGCGGCTACT 4450
Db 4569 CCACTAAGCAGAGATGAGAGAGGCGCCCAATTAAGCTCAACGCGGAGGCGGCTACT 4628
QY 4451 CTTTGAAGAGAGCAAGCTCATCCGCGCAGACATGACTTCAAAACCTTGTCTGAGCT 4510
Db 4629 CCTTGAAGAGAGCAAGCTCATCCGTCAGACATGAGTCAACAGACATGACCTTCACT 4688
QY 4511 GTCTGAGCCCAAGATGCAACAGCCCGAGGCTCCAGTAAAGATCTCAACTGTGANA 4570
Db 4689 GCGTGTGTCCGAGAGAGAGGAGCGCCAGGCTCCAGTAAAGATCTCAACTGTGANA 4748
QY 4571 CCACTGAGGCTCAAGAGAGAGATTCGTGATGCCATCTTCAAGAAATGTGCTGTGCC 4630
Db 4749 GCACTACCAAGGCAAGATTAAGTGTCTGACACTGTGTACAAAGGCACTTCCGTACTCC 4808
QY 4631 ACCGGCCAAAGCTGACAGATATGAGTGTGAGTGTGAGCAAGAGAGTGGGGCAAGATGA 4690
Db 4809 AGGCTCCAAAGCTGAGGACATGAGCTGAGTGTGAGGCGCCAGGCAATGACTGCGATCA 4868
QY 4691 TCTTGAAGATGAGAGATCAACCAAGATGAGAAATGATGAGAAAGGCACTGAACACAC 4750
Db 4869 TCTTGAAGATGAGAGATGAGAGATGAGAGATGAGAGATGAGAGATGAGAGATGAGAG 4928
QY 4751 TGGCCCACTACCAAGGCTGAGAGTGTGCTGAGTGTGAGTGTGAGTGTGAGTGTGAG 4810
Db 4929 TGGCCCACTACCAAGGCTGAGAGTGTGCTGAGTGTGAGTGTGAGTGTGAGTGTGAG 4988
QY 4811 CCACTAAGAGAGAGCAACTCAACGCTTCCAGAGCTCAAGCAAGTAAATTAATGAACA 4870
Db 4989 CCACTAAGAGAGAGCAACTCAACGCTTCCAGAGCTCAAGCAAGTAAATTAATGAACA 5045
QY 4871 TGAATCGGATCAAGGAGAGCCCGGACAGGCTCCGCTCAAGAGCACTTATGATCACTCTG 4930
Db 5046 TGTCTGCGCAAGGCAAGAGGCTGATAGGCTTCCGCTCAAGGAGCACTTATGATCACTCTG 5105
QY 4931 ACTTGAAGATGAGATCAAGATGTGCACTTGAAGAGAGAGAGAGAGAGAGAGAGAGAG 4990
Db 5106 ACCAGAGAGAGGAGCACTTATGAGAGCTTGTGAGAGAGAGAGAGAGAGAGAGAGAGAG 5165
QY 4991 AG 5050
Db 5166 GCGAGGAG 5225
QY 5051 CTAAAGGCACTGACAGAAAGTTTGTGATGAGCTCTTTGAGACCACTTCAAGAGAGGCAAC 5110

D	b		5226	CMAAGGGCACCTGCAGAAGTTGCGATGAAGACCTCTTGGAAACAAGTGTTCAAGCACAAGCCC	52859
O	y		5111	ACCCTGCGCTCGCCCTGCCCTTCAGCATCAAGTAATGTGGACTTCTGGATGAGCAGG	51707
D	b		5286	ACCAGGGCTCGGCCCTGCCCTGGCCAACAAAGTACATGTGTGAAGCTTCTGGATGAGCAGG	53454
O	y		5171	CTGATAAACATGAGCATTCATGACCCGCAGCTCCGCCTACCTGGAGAGCAATTGCGCTCG	52308
D	b		5346	CGGACCAAGCGCAGATCAGGCAACCCGATGTGGCCGACACCTGGAAAGCAACTGCGCTCG	54050
O	y		5221	CCCTGAGGTTTTGGGTCAACATGATCAAAGAACCCGAGTTGTGTTAACATCCATAGA	52939
D	b		5406	CGCTGCGCTTCTGGGTGAATGTGAACAAAGAACCCGAGTTGATGTTCAGCATCCACAAG	54655
O	y		5291	ACAGCATCACAGAAGCGCTCGCTCTGTGTGTGGCTCAAGACCTTCATGAGACTTGTCTCCA	53580
D	b		5466	ACAAGCATACCGAATGCTTGGCTGTGGTGTGATGCCBACCACTTCAAGAACCTCTGCTCTA	55253
O	y		5351	CGTCAGAGCACCGGTGGGCAAGGACCTGCCCCCTCCACAAGCTGCTGTATGCCAAGACA	54108
D	b		5526	CATCCGAGCACCGCTCGGGGAAAGCACTGCCCTCCAAAACCTGCTTACGCCAAGACCA	55859
O	y		5411	TCCCCAGGTACAAGATTGGGTGAGAGAGTTTATCTCAAGCATAGAGAAATGACAGCCA	54707
D	b		5586	TCCCCAACTAACAGAGCTGGGTGAGAGAGTATATCGAGACATTGCANAGATGGATGCCA	56454
O	y		5471	TCAGGACCAAGACATGAACGATACCTGGCTGAGCACTCCCGATGCACATGAATAGT	55303
D	b		5646	TCAGGAGCAAGAGACATGATGCTTACCTGGTAGACAGCTCCGCGCTCCACGCGCACGCACT	57058
O	y		5531	TCAACACCATGAGTCACTCTCAGAGATCTTCTCTATGTGGGCAAAATACAGCCGAGAGA	55908
D	b		5706	TCAGGCTCTGAGTGCCTCAACAGAGCTGTATTTTAATGTACCAAGTACCGCCAGAGAA	57663
O	y		5591	TCCTTGAACCTTGGAGCAACGATGACCAAGTGTGGAGACGAAMACTGGCTTCAAAACTAG	56555
D	b		5766	TTCTCAACGGCTCTGGACCGAGAGTGCCTCTTGTTCGAGAGGATTAAGTTGGCGGCAAGAACTGG	58225
O	y		5651	AACAAGTCATAACCTCAATGAGCTTAAACAGCTGA	5685
D	b		5826	AACGATATCAAGCTTCGTGTCTCAGCAGCAAGCTTAA	5860

RESULT 8
ABV76746
ID ABV76746 standard; cDNA; 3914 BP.

AC ABV76746;

DT 28-MAR-2003 (first entry)

Human plexin 2 protein 56.11-encoding cDNA

KM Human; plexin 2 protein 56.11; recombinant production; gene therapy
KM palsy; dementia; gene; ss.

OS Homo sapiens

EH	Key	Location/Qualifiers
FT	CDS	1391..2926
FT		/*tag= a
FT		/product= "Human plexin 2 protein 56.11"

PN CN1359955-A.

PD 24-JUL-2002

PF 20-DEC-2000; 2000CN-00135177.

PR 20-DEC-2000; 2000CN-00135177.

PA (BODE-) BODE GENE DEV CO LTD SHANGHAI.

XX Mao Y, Xie Y;
PI
XX
XX
DR WPI, 2002-733643/80.
DR P-P5D5; ABP58529.
XX
XX Polypeptide-human plexin 2 protein 56.11 and polynucleotide for coding
PT it.
PT
XX
XX
PS Claim 6; Page 24-26 (Disclosure); 35pp; Chinese.
PV

Claim 6; Page 24-26 (Disclosure); 35pp; Chinese.

The invention relates to human plexin 2 protein 56.11 (ABP985529) and nucleic acids encoding it (AB767646). The protein has a molecular weight of 56.11 kD. The invention also relates to a method for the recombinant production of the protein, an antagonist of the protein, and the use of the protein, gene and antagonist in therapeutic applications. Plexin 2 protein 56.11 can be used in the treatment of a variety of diseases such as palsy and dementia. The present sequence represents cDNA encoding human plexin 2 protein 56.11

Q Sequence 3914 BP; 924 A; 1181 C; 1016 G; 793 T; 0 U; 0 Other;

Query Match	17:18;	Score 974;	DB 6;	Length 3914;
-------------	--------	------------	-------	--------------

Matches 1527; Conservative 0; Mismatches 510; Indels 227; Gaps 2

QY	3023	CCATATCCTCAAGATGAGAGTCTCTAGAGATGAAAGGCTGCGATGACAGTGGACAGAGGCCAAGA	3082
Db	1416	CCCATATCTCCATGAGCGCTTGCGCCGGTCCCTCTTTCTGTGAGAGTGGACCGAGCCATG	1475
QY	3083	TCACACGAGACCTGGCTTTTCAAGTATGTGGAAGACCCCAACATCTGCGGATTTGACCCAG	3142
Db	1476	TGGATTAGCAACCTGACGATTTGAGTACATATATGACCTCGGGTTCACAGCATCAGCCAG	1535
QY	3143	AATGAGCATTTGTAGTGGAAACACACCCATCGCGATGAGGGAGCCACCTGAGACTCA	3202
Db	1536	AGTGAAGATTTGCGATGGCCACACACCCCTGACATCAAGGCTTAACTGTGAATGTCA	1595
QY	3203	TACAGAAACCCCGAGATCCGTGACCAAGCATGAGGGAAGAGACATCATATCTGTGAAG	3262
Db	1596	TTTCAGAGCCAAAGATCCGAGTCAAAATTCAATGGCAAAATCTGTCAATGTGTAAAG	1655
QY	3263	TTTCGAAACGTACTGAATATGACCTGTCAAGCGCCCGCCCTCGCTGAGGTCTTACACAC	3322
Db	1656	TTGTGAACACACACACCTTCACTCGCTGGACCTCTTGACACGAGCTAACCGCCCTG	1715
QY	3323	AGTCAGACCTTACCCGAGAGGCCGACAGAGATTGTGCTTCACTCTGAGACAAAGTCCAGTCC	3382
Db	1716	GCCGTGACACGTGTGGAACGCCCAATAGATTGTGATTTGTCTTAAACAATGTCCAAATCT	1775
QY	3383	TGCTGATCTCTCAACAAACCACTTACCTTACTATCTCCAAACCCGGGTGTTGAGGCTTTG	3442
Db	1776	TGCTAATTTTAAAGAACACCAAGTTTATCTAACCCCAACCCGACCTTTGAATCTGTTA	1835
QY	3443	GTCCCTCAGGAATCCTGGAAGCTCAAGCCTGGACAGCCCATCATCTTAAAGGGCAAGAAC	3502
Db	1836	GCCCTACTGAGAGTCTTGATATCAAAGACCAAGATGCCCATCATCTGTAGAGGGCAAAAAC	1895
QY	3503	TGATTCGCCCTGTGTGAGCTGGGGGCAACGTGAAGCTGAACCTACACTGTGTGTTGGAGGA	3562
Db	1896	TCTGGCCCTCTGCTCTGAGAG--GGCCAAACTCAACTACACTGTGCTCATCGAGAGA	1952
QY	3563	AGCGGTGACCGGTACACCGTGCATATGTCACAGCTGTGAGATGCCCCCAACCTCATCG	3622
Db	1953	CCCTTGCTGTCAACCGTATCTAGAACCAAGCTTCTCTGGAAGCTTCCAACTCACCG	2012
QY	3623	GCAGGCAAAAGTATGAGCCCGTGTGAGTGCATGAGTACTTCCCGGGAGTGTGTACA	3682
Db	2013	GCGAGCAAAAGTATGATGTCACGTGGGCGGGAGTGTGTTCTCCCTGTGCTCGGTGAGTG	2073
QY	3683	TTGCCCCGGAAGCCCGCTAGCCTTGACCCGCACTGTCAAGATGTGCAATGGCTGTGGCGCG	3742
Db	2073	TGATTTGAGAGCTTGTGTGACCTGTGCAAGCACTGTGAGATGTGAGATGTGCGCCGAGACCC	2132

QY 3743 TCTCATCATTTTTCATGCTGGCCCTGCTCATTTGCTTAAAGCAAGTCCCGGAAAGTG 3802
 Db 2113 TCTTCTCATCATGCTCATGCTGCTTCAATGCTTCAAGCCGCAAGTCTCGAAGAAATG 2192
 QY 3803 ACCTCAGCTGAGAGGCTGAGATGAGATGAGCAACCTGGAATCCCGTGTGGCCCTGG 3862
 Db 2193 ACTTCACCTCTCAAGGGCTGGAATGAGATGAGCAATCTGGAATCCCGTGTGGCCCTGG 2252
 QY 3863 AGTGAAGAAAGCTTTTCCGAGCTGAGAGAGCAATCATGAGCTGACAGATGACTGG 3922
 Db 2253 AGTGAAGAAAGCTTTTCTGAGCTCCAGAGAGATATCAATGATGACAGATGACTGG 2312
 QY 3923 ATGAGAGCGGGATTCGTTCCGAGCTATGAACTTACACCATGCGGGTGTGTTCCAG 3982
 Db 2313 ACCGCTCAGAAATCCCTTAACCTGAGCTTCTGACTGAGCTATGCGAGTCTGTTCCCG 2372
 QY 3983 GAATTGAAGACCACTGCTCTCCGAGACCTTGAGGTCCTCGGCTACCGGCAAGAGCTGG 4042
 Db 2373 GCATGAGAGACCACTCCCTCTCGCGAGAGCTGAGAGTACAAAGAAAGCGGCAAGACAG 2432
 QY 4043 TGGAAAGAGCTGAGAGCTTCCGCGAGCTCATCAACAAGAGTGTCTGCTGCTCT 4102
 Db 2433 TGGAAAGAGCTTGAAGCTCTTGGCCAGCTCATCAACAAGAGTGTCTGCTGCTCT 2492
 QY 4103 TCATCCGACGCTTGAAGTCCAGAGCTTCTCCATGCGGAGCCGTCGCAAGCTGGCT 4162
 Db 2493 TCATCCGACGCTTGAAGTCCAGAGCTTCTCCATGCGGAGCCGTCGCAAGCTGGCT 2552
 QY 4163 CACTCATATGACCTGCTGTCAGAGCAAGCTGAGTACGCACTGATGCTGTAAGAGC 4222
 Db 2553 CGCTCATATGACCTGCTGTCAGAGCCGCTGAGATATGCACTGATGCTTCAAGAGC 2612
 QY 4223 TCTGAGCGAGCTCATTTGACAAAGCTGAGAGAGCAAGCAACCTTAAGCTGCTGCTCA 4282
 Db 2613 TCTGCTCTGAGCTCATGCTGATGAACTGAGAAACAAGCAACCTTAAGCTGCTGCTC 2672
 QY 4283 GAGAGCTGAGTCACTGCTGTCAGAGCAAGCTGAGTACGCACTGATGCTGTAAGAGC 4342
 Db 2673 GAGAGCTGAGTCACTGCTGTCAGAGCAAGCTGAGTACGCACTGATGCTGCTGCTCA 2732
 QY 4343 AGTTCCTCA----- 4351
 Db 2733 AGTTCCTCAAGTGAAGAAAGGAGTGAAGAGCGGCGCTTACGCTCAAGAGAGATGGA 2792
 QY 4352 ----- 4351
 Db 2793 GCTGCTAGATCTGCGCACAGGCTGATGAGGACCCCTTCCCTGAGAGCAAGAACCT 2852
 QY 4352 ----- 4351
 Db 2853 GGAAGTGGCTGTGCTAAAGCCCTTCTTCACTGGGCTCGCTCCGAGAGAAAGGAGAA 2912
 QY 4352 -----AGGAGTG 4358
 Db 2913 ACGTACAACTGAGAGAGCGCTGACCACACTCTGCTGCTGCTGCTGCTGCTGCTGCTG 2972
 QY 4359 TCTGCTGAGAGCCCTTCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 4418
 Db 2973 CGAGAGGAGCCACTTCTTCACTGCTTACTGCTGCTGCTGCTGCTGCTGCTGCTGCT 3032
 QY 4419 CATTTGAGCATCAAGGAGAGCGCGCTTCTTCTGAGCGAGCAAGGCTCATCTGCGCA 4478
 Db 3033 CATTTGAGCATCAAGGAGAGCGCGCTTCTTCTGAGCGAGCAAGGCTCATCTGCGCA 3092
 QY 4479 GCAAGATTGACTAAGAAAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 4538
 Db 3093 GCAAGATTGACTAAGAAAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 3152
 QY 4539 CGAGGCTCCAGTAAAGATCTCAACTGTGACACATCACTCAAGGCTCAAGAGAGATTTCT 4598
 Db 3153 AGAGATCCCAAGTAAAGTGTAAAGTGTGACACATCACTCAAGGCTCAAGAGAGATTTCT 3212

QY 4599 GGATGCATCTTCAAGAAATGTCCTTGTCTCCACCGGCCCAAGCTGAGATATGATCT 4658
 Db 3213 TGAATCCGCTGTAAGAAATGTCCTTATTCAGAGGCGGAGGAGATGATGATCT 3272
 QY 4659 GGAATGGCGCAAGAAAGTGGGGCAAGATGATCTTGAAGATGAAGCATCAACCA 4718
 Db 3273 GGAATGGCGCGCAAGGCGCGAGATGCGCGGATGCTGTAAGATGAGCATCAACCA 3332
 QY 4719 GATTGAGATGATTGAAGAGCACTGAACACACTGGCCCACTAACAGATGATGTTT 4778
 Db 3333 GATTGAGATGATTGAAGAGCACTGAACACACTGATGATGATGATGATGATGATGAT 3392
 QY 4779 CGTGTGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 4838
 Db 3393 GGTGTGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 3452
 QY 4839 CTCCAGAGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 4898
 Db 3453 CTCCAGAGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 3512
 QY 4899 CCTCCGCTCAGAGACACTATGATCACTTCTGAGAGTGAAGTGAATGATGATGAT 4958
 Db 3513 CCTCCGCTCAGAGACACTATGATCACTTCTGAGAGTGAAGTGAATGATGATGAT 3572
 QY 4959 CCTAGTGAAGAACCAAGACAGAGACAGAGACAGAGAGAGAGAGAGAGAGAGAGAG 5018
 Db 3573 TCTGTGAAGAACCAAGACAGAGACAGAGACAGAGAGAGAGAGAGAGAGAGAGAG 3632
 QY 5019 GTCTGAATCTACTGATCCGAGCTCTGAGCACTTCTGAGCACTTGAAGGCAAGC 5062
 Db 3633 GTCTGAATCTACTGATCCGAGCTCTGAGCACTTCTGAGCACTTGAAGGCAAGC 3676

RESULT 9
 ABK70006
 ID ABK70006 standard; DNA; 2597 BP.
 AC ABK70006;
 XX
 DT 15-JUL-2002 (first entry)
 XX
 DE cDNA encoding human Pro peptide #46.
 XX
 KM Human; ss; gene; PRO; secreted protein; transmembrane protein;
 KM genetic disorder; tumor; cancer.
 OS Homo sapiens.
 XX
 PN WO20022488-A2.
 PD 28-MAR-2002.
 PF 29-AUG-2001; 2001MO-US027099.
 XX
 ER 01-SEP-2000; 2000US-0229896P.
 ER 05-SEP-2000; 2000US-0230621P.
 ER 22-SEP-2000; 2000US-0235147P.
 ER 10-NOV-2000; 2000MO-US030873.
 ER 12-JAN-2001; 2001US-0261878P.
 ER 16-JAN-2001; 2001US-0261910P.
 ER 16-JAN-2001; 2001US-0261939P.
 ER 16-JAN-2001; 2001US-0262150P.
 ER 25-JAN-2001; 2001US-0264395P.
 ER 02-FEB-2001; 2001US-0264421P.
 ER 09-FEB-2001; 2001US-0267623P.
 ER 28-FEB-2001; 2001MO-US006520.
 ER 09-MAR-2001; 2001US-0274399P.
 ER 03-APR-2001; 2001US-0280982P.
 ER 04-APR-2001; 2001US-0282129P.
 ER 04-APR-2001; 2001US-0282199P.
 ER 09-MAY-2001; 2001US-0290589P.
 ER 25-MAY-2001; 2001MO-US017092.
 ER 01-JUN-2001; 2001MO-US017800.

PR 20-JUN-2001; 2001MO-US019692.
 PR 29-JUN-2001; 2001MO-US021066.
 PR 09-JUL-2001; 2001MO-US021735.
 XX
 PA (GENTH) GENENTECH INC.
 PI Baker KP, Eaton DL, Filvaroff E, Goddard A, Grimaldi JC,
 PI Gurney AL, Smith V, Stephan J, Watanabe CK, Wood WI, Zhang Z,
 PI Fong S;
 XX
 DR WPI: 2002-362426/39.
 DR P-PSDB; ABG34075.
 XX
 PT New PRO polypeptides and polynucleotides encoding the polypeptides,
 PT useful in gene therapy, chromosome identification, tissue typing, or for
 PT genetic analysis of individuals with genetic disorders.
 XX
 PS Claim 2; Fig 91; 218pp; English.
 CC This invention relates to the cDNA and protein sequences of novel
 CC secreted and transmembrane polypeptides PRO polypeptides. The invention
 CC also comprises a method for producing the proteins of the invention by
 CC recombinant means and antibodies specific for the protein of the
 CC invention. The antibody may be used for detecting the PRO proteins of the
 CC invention and may be used to modify their activity. polynucleotides may
 CC be used as hybridisation probes for a cDNA library to isolate the full-
 CC length PRO cDNA or to isolate other cDNAs, to construct hybridisation
 CC probes for mapping the gene which encodes that PRO and for genetic
 CC analysis of individuals with genetic disorders, in assays to identify
 CC other proteins or molecules involved in binding reaction, to generate
 CC transgenic animals or knock-out animals which in turn are useful in the
 CC development and screening of therapeutically useful reagents, for
 CC chromosome identification, and tissue typing. The PRO polypeptides are
 CC useful in gene therapy, and as molecular weight markers for protein
 CC electrophoresis purposes. The sequences may also be used to detect
 CC overexpression on PRO polypeptides in cancerous tumours and for screening
 CC for differentially expressed genes using microarray technology. The
 CC present sequence represents a cDNA encoding a human PRO protein of the
 CC invention
 CC
 XX
 SQ Sequence 2597 BP; 688 A; 570 C; 642 G; 697 T; 0 U; 0 Other;
 Query Match 13.9%; Score 793.8; DB 6; Length 2597;
 Best Local Similarity 99.7%; Pred. No. 6; 1e-181;
 Matches 799; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
 QY 579 CGAGTATTTTCCCACTCTCCAGCCGGAACAGCAAGAACTCTGAGCGGATGGCAT 638
 DB 1 CGAGTATTTTCCCACTCTCCAGCCGGAACAGCAAGAACTCTGAGCGGATGGCAT 60
 QY 639 GTTCGCGTACGTCCTTCATGATGATGTCGTGCGCTCGATGATTAGATCCTTCGACAC 698
 DB 61 GTTCGCGTACGTCCTTCATGATGATGTCGTGCGCTCGATGATTAGATCCTTCGACAC 120
 QY 699 CTTCACCATCATCCCTGACCTTTGATATCTATGTCATGAGCTTTTGAAGTGGCAACTT 758
 DB 121 CTTCACCATCATCCCTGACCTTTGATATCTATGTCATGAGCTTTTGAAGTGGCAACTT 180
 QY 759 TGTCTACTTTTGAACCTTCCAACTGAGATGATGTCCTCCACAGGCTCCACCAAGGA 818
 DB 181 TGTCTACTTTTGAACCTTCCAACTGAGATGATGTCCTCCACAGGCTCCACCAAGGA 240
 QY 819 GCAGGTGATATACCAAGCTCGTGAAGCTTTTGAAGAGAGACACAGCTTCACCTCTA 878
 DB 241 GCAGGTGATATACCAAGCTCGTGAAGCTTTTGAAGAGAGACACAGCTTCACCTCTA 300
 QY 879 TGTAGAGTGGCCATTGGCTGTGAGCAGTGGGGTGAAGTACCGGCTGTGAGGCTGC 938
 DB 301 TGTAGAGTGGCCATTGGCTGTGAGCAGTGGGGTGAAGTACCGGCTGTGAGGCTGC 360
 QY 939 CTACCTGTCCAAAGCGGGGGCGGTGCTTGGCAGAGACCTTTGAGTCCATCCAGATATGA 998
 DB 361 CTACCTGTCCAAAGCGGGGGCGGTGCTTGGCAGAGACCTTTGAGTCCATCCAGATATGA 420

QY 999 CTGCTCTTTCACCGTCTTCTTCCAAAGGCGCAAGACGGAAATGAAATCCTTGATGATC 1058
 DB 421 CCTGCTCTTTCACCGTCTTCTTCCAAAGGCGCAAGACGGAAATGAAATCCTTGATGATC 480
 QY 1059 GGCCCTGTCATCTTCACTCTTGAAGCAGATAAATGACCGGATTAGGAGCGGCTGCAGTC 1118
 DB 481 GGCCCTGTCATCTTCACTCTTGAAGCAGATAAATGACCGGATTAGGAGCGGCTGCAGTC 540
 QY 1119 TTGTTACCGGGGCGAGGGCAGCGCTGACCTGCGCTGCTCAAGGTGAAGACATCCCTG 1178
 DB 541 TTGTTACCGGGGCGAGGGCAGCGCTGACCTGCGCTGCTCAAGGTGAAGACATCCCTG 600
 QY 1179 CAGCAGTGGCTCTTACCATTAACCATTAACCATTAACCATTAACCATTAACCATTAACCAT 1238
 DB 601 CAGCAGTGGCTCTTACCATTAACCATTAACCATTAACCATTAACCATTAACCATTAACCAT 660
 QY 1239 GGGAGTGTCCGACATGTCGTGTCGTGTAATCCGCTTTCACGAGACAGGACCGCATGAC 1298
 DB 661 GGGAGTGTCCGACATGTCGTGTCGTGTAATCCGCTTTCACGAGACAGGACCGCATGAC 720
 QY 1299 GTCTGTATCGCATATGTCCTTAACCAAGACCACTCTGCTGCTTGTGGCCACCAAGTGG 1358
 DB 721 GTCTGTATCGCATATGTCCTTAACCAAGACCACTCTGCTGCTTGTGGCCACCAAGTGG 780
 QY 1359 CAAGTGAAGAAGATCC 1375
 DB 781 CAAGTGAAGAAGATCC 797
 XX
 RESULT 10
 ID ADA01359
 ADADA01359 standard; cDNA; 2597 BP.
 XX
 AC ADA01359;
 XX
 DT 06-NOV-2003 (first entry)
 XX
 DE Human PRO polynucleotide #46.
 XX
 KW Human; gene; ss; PRO: secreted polypeptide; transmembrane polypeptide;
 KW tumor necrosis factor-alpha; TNF-alpha; blood; chondrocyte cell; tumour;
 KW adrenal; lung; colon; breast; prostate; rectum; cervix; liver; cancer;
 KW microvascular endothelial cell; endothelial cell tube formation;
 KW sports-related joint problem; articular cartilage defect; osteoarthritis;
 KW rheumatoid arthritis; osteopathic; antirheumatic; antiarthritic.
 XX
 OS Homo sapiens.
 XX
 EN US2003068779-A1.
 XX
 PD 10-APR-2003.
 XX
 PF 16-SEP-2002; 2002US-00245107.
 XX
 PR 09-MAY-2001; 2001US-0290589P.
 PR 29-AUG-2001; 2001MO-US027099.
 PR 18-JUL-2002; 2002US-00197942.
 XX
 PA (GENTH) GENENTECH INC.
 PI Baker KP, Eaton DL, Filvaroff E, Goddard A, Grimaldi JC,
 PI Gurney AL, Smith V, Stephan J, Watanabe CK, Wood WI, Zhang Z,
 PI Fong S;
 XX
 DR WPI: 2003-625484/59.
 DR P-PSDB; ADA01360.
 XX
 PT Novel isolated PRO1313, PRO20080 or PRO21383 polypeptide useful for
 PT stimulating proliferation of human microvascular endothelial cells, and
 PT PRO6018 polypeptide useful for stimulating proliferation of chondrocyte
 PT cells.

PS Claim 2; Fig 91; 307bp; English.

XX The invention relates to isolated human PRO polypeptides (secreted and
CC transmembrane polypeptides) and the polynucleotides encoding them. The
CC invention also relates to an antibody which specifically binds to a PRO
CC polypeptide, a method for stimulating the release of tumour necrosis
CC factor-alpha (TNF-alpha) from human blood, a method for stimulating the
CC proliferation or differentiation of chondrocyte cells and a method for
CC detecting the presence of a tumour in a mammal (e.g. adrenal, lung,
CC colon, breast, prostate, rectal, cervical and liver tumours). The
CC polynucleotides are useful in molecular biology, including uses as
CC hybridisation probes, in chromosome and gene mapping, in generating
CC antisense RNA and DNA and in gene therapy. The polynucleotides may also
CC be used in preparing PRO polypeptides by recombinant techniques and in
CC generating either transgenic animals or knock-out animals which are
CC useful in the development and screening of therapeutically useful
CC reagents. The PRO polypeptides or antibodies are used in preparing a
CC medicament for treating a condition responsive to the polypeptides or
CC antibodies, such as tumours, for stimulating and inhibiting proliferation
CC of human microvascular endothelial cells, for inducing endothelial cell
CC tube formation and for treating sports-related joint problems, articular
CC cartilage defects, osteoarthritis and rheumatoid arthritis. This sequence
CC represents a human PRO polynucleotide of the invention.

SQ Sequence 2597 BP; 688 A; 570 C; 642 G; 697 T; 0 U; 0 Other;

Query Match 13.9%; Score 793.8; DB 8; Length 2597;
Best Local Similarity 99.7%; Pred. No. 6,1e-181;

Matches 795; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 579 CGAGTATTTTCCACCATCTCCAGCGGAACTGACCAAGAACTCTGAGCGGATGGCAT 638
DB 1 CGAGTATTTTCCACCATCTCCAGCGGAACTGACCAAGAACTCTGAGCGGATGGCAT 60
QY 639 GTTCGGGTACGCTTTCATGATGAGTTCGTGGCTCGATGATTAAGATCCCTCGGACAC 698
DB 61 GTTCGGGTACGCTTTCATGATGAGTTCGTGGCTCGATGATTAAGATCCCTCGGACAC 120
QY 699 GTTCACCATCATCCTGACTTGAATGATGATGATGATGATGATGATGATGATGATGAT 758
DB 121 GTTCACCATCATCCTGACTTGAATGATGATGATGATGATGATGATGATGATGATGAT 180
QY 759 TGTCTACTTTTGAACCTCCAACTGAGATGAGTGTCCACAGGCTCCACCAAGAA 818
DB 181 TGTCTACTTTTGAACCTCCAACTGAGATGAGTGTGTCTCCACAGGCTCCACCAAGAA 240
QY 819 GCGAGTGTATACATCCAGCTCGTGAAGGCTTTGCAAGAGACACAGCTTCACTCTTA 878
DB 241 GCGAGTGTATACATCCAGCTCGTGAAGGCTTTGCAAGAGAGACACAGCTTCACTCTTA 300
QY 879 TGTAGAGGTGCCCATGCTGAGAGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 938
DB 301 TGTAGAGGTGCCCATGCTGAGAGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 360
QY 939 CTACCTGTCTCCAAAGCGGGGCGCTGCTTGGAGAGAGAGAGAGAGAGAGAGAGAGAG 998
DB 361 CTACCTGTCTCCAAAGCGGGGCGCTGCTTGGAGAGAGAGAGAGAGAGAGAGAGAGAG 420
QY 999 CCGGCTCTTCAACGCTTCTCCAGAGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1058
DB 421 CCGGCTCTTCAACGCTTCTCCAGAGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 480
QY 1059 GGGCTGTGATCTTTCATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1118
DB 481 GGGCTGTGATCTTTCATGATGATGATGATGATGATGATGATGATGATGATGATGAT 540
QY 1119 TTTTTCACCGGGGAG 1178
DB 541 TTTTTCACCGGGGAG 600
QY 1179 CAGCAGGTGCGCTCTTAACATGATGATGATGATGATGATGATGATGATGATGATGAT 1238
DB 601 CAGCAGGTGCGCTCTTAACATGATGATGATGATGATGATGATGATGATGATGATGAT 660

QY 1239 GGGAGTGTCCGACATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1298
DB 661 GGGAGTGTCCGACATGATGATGATGATGATGATGATGATGATGATGATGATGAT 720
QY 1299 GTCTGTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1358
DB 721 GTCTGTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 780
QY 1359 CAAGCTGGAAGAGATCC 1375
DB 781 CAAGCTGGAAGAGATGTC 797

RESULT 11

ID ADA43788 standard; cDNA; 2597 BP.
XX ADA43788;
AC ADA43788;
XX 20-NOV-2003 (first entry)
DT 20-NOV-2003 (first entry)
XX Human cDNA encoding secreted/transmembrane polypeptide PRO34003.

XX ss; gene; human; PRO; secreted protein; transmembrane protein;
KM endothelial cell tube formation; chondrocyte cell differentiation;
KM microvascular endothelial cell; tumour; lung tumour; colon tumour;
KM breast tumour; prostate tumour; rectal tumour; kidney tumour;
KM liver tumour; cytostatic; vaccine.

XX Hemo sapiens.

XX US2003064474-A1.

XX 03-APR-2003.

XX 16-SEP-2002; 2002US-00245859.

XX 29-AUG-2001; 2001WO-US027059.

XX 18-JUL-2002; 2002US-00197942.

XX (GENTH) GENENTECH INC.

XX Baker KP, Baton DU, Filvaroff B, Goddard A, Grimaldi JC;
PI Guney AL, Smith V, Stephan JP, Watanabe CK, Wood WI, Zhang Z;
PI Fong S;

XX WPI; 2003-605867/57.

XX P-PSDB; ADA43789.

XX New isolated nucleic acid encoding a PRO polypeptide, e.g. PRO20080 or
PT PRO2183, useful in molecular biology, chromosome and gene mapping, in
PT generating antisense RNA and DNA, and in gene therapy.

XX Claim 2; Fig 91; 308bp; English.

XX The invention relates to an isolated secreted/transmembrane (PRO)
CC polypeptide, having at least 80% sequence identity to a sequence selected
CC from any one of the 57 amino acid sequences given in specification, or to
CC a sequence encoded by a nucleic acid molecule selected from any one of
CC the nucleic acids deposited under any of the ATCC accession numbers given
CC in specification, or a sequence having at least 80% identity to PRO
CC lacking its associated signal peptide, an extracellular domain of PRO
CC with or without its associated signal peptide. Also included are vectors,
CC transformed host cells, anti-PRO antibodies, the nucleic acids encoding
CC PRO, PRO fusion proteins, inducing endothelial cell tube formation (by
CC administering PRO281, PRO1560, PRO189, PRO4499, PRO6308, PRO6000,
CC PRO10275, PRO21207, PRO20933 or PRO34274 polypeptide or its agonist) and
CC an oligonucleotide probe derived from any one of the above nucleotide
CC sequences. PRO6018 polypeptide is useful for stimulating the
CC proliferation or differentiation of chondrocyte cells. PRO1313, PRO20080
CC and PRO2183 polypeptides are useful for stimulating the proliferation of
CC human microvascular endothelial cells. PRO6071, PRO4487 and PRO6006

CC polypeptides are useful for inhibiting the proliferation of human
 CC microvascular endothelial cells. PRO polypeptides are useful for
 CC detecting the presence of tumour in a mammal, including tumours of lung,
 CC colon, breast, prostate, rectal, kidney and liver. PRO281, PRO1560,
 CC PRO189, PRO499, PRO6308, PRO10275, PRO1207, PRO20933 and
 CC PRO34274 polypeptides are useful for inducing endothelial cell tube
 CC formation. PRO or the antibody are useful in the preparation of a
 CC medication for treating a condition responsive to PRO polypeptide. The
 CC oligonucleotide probes are useful for isolating genomic and cDNA
 CC nucleotide sequences, for measuring or detecting the expression of an
 CC associated gene, and as antisense probes. PRO nucleic acid is useful as a
 CC hybridisation probe, in chromosome and gene mapping, in the generation of
 CC antisense RNA and DNA, and for the preparation PRO polypeptides. PRO and
 CC PRO nucleic acid are useful as therapeutic agents, e.g. vaccines. The
 CC present sequence encodes a PRO protein.

XX
 XX Sequence 2597 BP; 688 A; 570 C; 642 G; 697 T; 0 U; 0 Other;
 SQ

Query Match 13.9%; Score 793.8; DB 8; Length 2597;
 Best Local Similarity 99.7%; Pred. No. 6.1e-181;
 Matches 795; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 579 CGAGTATTTTCCACCATCTCCGACCGGAACTGACAGAACTCTGAGCGGATGCGAT 638
 DB 1 CGAGTATTTTCCACCATCTCCGACCGGAACTGACAGAACTCTGAGCGGATGCGAT 60
 QY 639 GTTCGCGTACGCTTCCATGATGATGATGATGATGATGATGATGATGATGATGATGAT 698
 DB 61 GTTCGCGTACGCTTCCATGATGATGATGATGATGATGATGATGATGATGATGATGAT 120
 QY 699 CTTACCATCATCTCCGATCTTATATCTATCTATCTATCTATCTATCTATCTATCTATCT 758
 DB 121 CTTACCATCATCTCCGATCTTATATCTATCTATCTATCTATCTATCTATCTATCTATCT 180
 QY 759 TGTCTACTTTTACCCCTCCACCTGAGATGATGATGATGATGATGATGATGATGATGATGAT 818
 DB 181 TGTCTACTTTTACCCCTCCACCTGAGATGATGATGATGATGATGATGATGATGATGATGAT 240
 QY 819 GCAGGTGTATATCAACAGCTGCTGAGGCTTTGCAAGAGACAGACCTTCACTCTCTA 878
 DB 241 GCAGGTGTATATCAACAGCTGCTGAGGCTTTGCAAGAGAGACAGACCTTCACTCTCTA 300
 QY 879 TGTNAGAGTGGCCATGCTGCTGAGGCGATGAGGCGATGAGGCGATGAGGCGATGAGGCTGC 938
 DB 301 TGTNAGAGTGGCCATGCTGCTGAGGCGATGAGGCGATGAGGCGATGAGGCGATGAGGCTGC 360
 QY 939 CTAACCTGTCCAAAGCGGGGCGCTGCTGAGCAGACCTTTGAGTCCATCCAGATGATGA 998
 DB 361 CTAACCTGTCCAAAGCGGGGCGCTGCTGAGCAGACCTTTGAGTCCATCCAGATGATGA 420
 QY 999 CCGTCTCTTACCGCTCTTCTCCAAAGGCGGCAAGCGGAAATGAAATCCCTGGATGATGTC 1058
 DB 421 CCGTCTCTTACCGCTCTTCTCCAAAGGCGGCAAGCGGAAATGAAATCCCTGGATGATGTC 480
 QY 1059 GAGCCTGTGATCTTCACTTTGAGAGCAGATGAATGACCGCATTTAGAGCGGCTGAGTGC 1118
 DB 481 GAGCCTGTGATCTTCACTTTGAGAGCAGATGAATGACCGCATTTAGAGCGGCTGAGTGC 540
 QY 1119 TTGTATCCGGGGGCGAGGCGACCTGAGACCTGCTGAGCTTAAGTGAAGACATCCCTCG 1178
 DB 541 TTGTATCCGGGGGCGAGGCGACCTGAGACCTGCTGAGCTTAAGTGAAGACATCCCTCG 600
 QY 1179 CACGAGTGGCGCTCTTAACATGACATTAATCTGCTGAGCTGAGATGATGATGATGATGAT 1238
 DB 601 CACGAGTGGCGCTCTTAACATGACATTAATCTGCTGAGCTGAGATGATGATGATGATGAT 660
 QY 1239 GGGAGTGTCCGACATGTGTGCTGATTTCCCGCTTTCACGAGAGACAGGAGCCGATGATC 1298
 DB 661 GGGAGTGTCCGACATGTGTGCTGATTTCCCGCTTTCACGAGAGACAGGAGCCGATGATC 720
 QY 1299 GTCTGTATGCAATATGTCTACAAAGACACTCTTGCTTGCTGGGACCAAAAGTGG 1358
 DB 721 GTCTGTATGCAATATGTCTACAAAGACACTCTTGCTTGCTGGGACCAAAAGTGG 780

QY 1359 CAACTGAGAGAGATCC 1375
 DB 781 CAACTGAGAGAGATGC 797

RESULT 12
 ADA43556
 ID ADA43556 standard; cDNA; 2597 BP.
 AC ADA43556;
 XX
 XX 20-NOV-2003 (first entry)
 DT
 XX
 DE Human cDNA encoding secreted/transmembrane polypeptide PRO34003.

XX
 KW ss; gene; human; PRO; secreted protein; transmembrane protein;
 KW endothelial cell tube formation; chondrocyte cell differentiation;
 KW microvascular endothelial cell; tumour; lung tumour; colon tumour;
 KW breast tumour; prostate tumour; rectal tumour; kidney tumour;
 KW liver tumour; cytostatic; vaccine.

XX
 OS Homo sapiens.
 XX
 PN US2003073196-A1.
 XX
 PD 17-APR-2003.
 XX
 PF 18-SEP-2002; 2002US-00246210.
 XX
 PR 04-APR-2001; 2001US-0282199P.
 PR 29-AUG-2001; 2001WO-US027099.
 FR 18-JUL-2002; 2002US-00197942.

XX
 PA (GENTH) GENENTECH INC.
 XX
 PI Baker KP, Eaton DV, Filvaroff E, Goddard A, Grimaldi JC,
 PI Gurney AL, Smith V, Stephan JP, Watanabe CX, Wood WI, Zhang Z,
 PI Fong S;
 XX
 DR WPI; 2003-743814/70.
 DR P-P5DB; ADA43557.

XX
 PT New secreted and transmembrane PRO polypeptides, e.g. PRO20080 or
 PT PRO3183 useful for stimulating the proliferation or differentiation of
 PT chondrocyte cells and detecting the presence of a tumor in a mammal.

XX
 PS Claim 2; Fig 91; 307pp; English.

XX
 CC The invention relates to an isolated secreted/transmembrane (PRO)
 CC polypeptide, having at least 80% sequence identity to a sequence selected
 CC from any one of the 57 amino acid sequences given in specification, or to
 CC a sequence encoded by a nucleic acid molecule selected from any one of
 CC the nucleic acid deposited under any of the ATCC accession numbers given
 CC in specification, or a sequence having at least 80% identity to PRO
 CC lacking its associated signal peptide, an extracellular domain of PRO
 CC with or without its associated signal peptide. Also included are vectors,
 CC transformed host cells, anti-PRO antibodies, the nucleic acids encoding
 CC PRO, PRO fusion proteins, inducing endothelial cell tube formation (by
 CC administering PRO281, PRO189, PRO499, PRO6308, PRO10275, PRO1207,
 CC PRO20933 or PRO34274 polypeptide or its agonist) and
 CC an oligonucleotide probe derived from any one of the above nucleotide
 CC sequences. PRO6018 polypeptide is useful for stimulating the
 CC proliferation or differentiation of chondrocyte cells. PRO313, PRO20080
 CC and PRO3183 polypeptides are useful for stimulating the proliferation of
 CC human microvascular endothelial cells. PRO6071, PRO487 and PRO6006
 CC polypeptides are useful for inhibiting the proliferation of human
 CC microvascular endothelial cells. PRO polypeptides are useful for
 CC detecting the presence of tumour in a mammal, including tumours of lung,
 CC colon, breast, prostate, rectal, kidney and liver. PRO281, PRO1560,
 CC PRO189, PRO499, PRO6308, PRO10275, PRO1207, PRO20933 and
 CC PRO34274 polypeptides are useful for inducing endothelial cell tube
 CC formation. PRO or the antibody are useful in the preparation of a

medicament for treating a condition responsive to PRO polypeptide. The CC oligonucleotide probes are useful for isolating genomic and cDNA CC nucleotide sequences, for measuring or detecting the expression of an CC associated gene, and as antisense probes. PRO nucleic acid is useful as a CC hybridisation probe, in chromosome and gene mapping, in the generation of CC antisense RNA and DNA, and for the preparation PRO polypeptides. PRO and CC PRO nucleic acid are useful as therapeutic agents, e.g. vaccines. The CC present sequence encodes a PRO protein.

XX Sequence 2597 BP; 688 A; 570 C; 642 G; 697 T; 0 U; 0 Other;

Query Match 13.9%; Score 793.8; DB 8; Length 2597;

Best Local Similarity 99.7%; Pred. No. 6,1e-181; Mismatches 2; Indels 0; Gaps 0;

579 CGAGTATTTTCCACCATCTCCAGCCGGAAGTACCAAGAACTCTGAGCGGATGGCAT 638
1 CGAGTATTTTCCACCATCTCCAGCCGGAAGTACCAAGAACTCTGAGCGGATGGCAT 60
639 GTTCGGTACGCTCTTCATGATGAGTTCGTGGCTCGATGATTAAGATCCCTTGGACAC 698
61 GTTCGGTACGCTCTTCATGATGAGTTCGTGGCTCGATGATTAAGATCCCTTGGACAC 120
699 GTTCACCATCATCTCCGATCTTGATATCTACTATGTTTGGTTTACAGTGGCAATT 758
121 GTTCACCATCATCTCCGATCTTGATATCTACTATGTTTGGTTTACAGTGGCAATT 180
759 TGTCTACTTTTGAACCTTCAACCTGAGATGTTGTTCTCCACAGGCTCCACACCAAGA 818
181 TGTCTACTTTTGAACCTTCAACCTGAGATGTTGTTCTCCACAGGCTCCACACCAAGA 240
819 GCAAGTGTATACATCCAACTGTTGAGGCTTTGCAAGAGGACAAAGCTTCAATCTTA 878
241 GCAAGTGTATACATCCAACTGTTGAGGCTTTGCAAGAGGACAAAGCTTCAATCTTA 300
879 TGTAGAGTGTCCCATGAGTGTGAGGAGGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTG 938
301 TGTAGAGTGTCCCATGAGTGTGAGGAGGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTG 360
939 CTACCTGTCCAAAGCGGGGCGGTGTTGGAGAGACCTTGGATTCATCCAGATGATGA 998
361 CTACCTGTCCAAAGCGGGGCGGTGTTGGAGAGACCTTGGATTCATCCAGATGATGA 420
999 CCGTCTCTTCAACCGCTTCTCCAAAGGGGCGGAAATGAAATCCCTGGATGATGTC 1058
421 CCGTCTCTTCAACCGCTTCTCTCCAAAGGGGCGGAAATGAAATCCCTGGATGATGTC 480
1059 GGGCCGTGTGATCTTTCATCTTGAAGAGATTAATGACCGCATTAAGAGCGGCTGCAATC 1118
481 GGGCCGTGTGATCTTTCATCTTGAAGAGATTAATGACCGCATTAAGAGCGGCTGCAATC 540
1119 TTGTTTCCGGGGGAGAGGAGCGTGAAGCTGGCTGGCTCAAGGGTGAAGAGATCCCTG 1178
541 TTGTTTCCGGGGGAGAGGAGCGTGAAGCTGGCTGGCTCAAGGGTGAAGAGATCCCTG 600
1179 CACAGTGTGCGCTCTTAACATTAAGATTAATGACCGCATTAAGAGCGGCTGCAATC 1238
601 CACAGTGTGCGCTCTTAACATTAAGATTAATGACCGCATTAAGAGCGGCTGCAATC 660
1239 GGAAGTGTCCGAGATGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTG 1298
661 GGAAGTGTCCGAGATGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTG 720
1299 GTCTGTATGCAATATGTCTTACAAGAACCACTCTCTGCTTGTGTGTTGTTGTTGTTG 1358
721 GTCTGTATGCAATATGTCTTACAAGAACCACTCTCTGCTTGTGTGTTGTTGTTGTTG 780
1359 CAAGCTGAAGAAGATCC 1375
781 CAAGCTGAAGAAGATCC 797

RESULT 13

ADA01231
ID ADA01231 standard; cDNA; 2597 BP.

XX ADA01231;

XX 06-NOV-2003 (first entry)

DE Human PRO polynucleotide #46.

XX Human, gene; se: PRO; secreted polypeptide; transmembrane polypeptide;

KM tumour necrosis factor- α ; TNF- α ; blood; chondrocyte cell; tumour;

KM cancer; lung; colon; breast; prostate; rectum; kidney; liver;

XX microvascular endothelial cell; endothelial cell tube formation.

OS Homo sapiens.

XX US2003068782-A1.

XX 10-APR-2003.

XX 16-SEP-2002; 2002US-00245851.

XX 27-APR-1999; 99US-0131271P.

XX 29-OCT-1999; 99US-0162506P.

XX 02-DEC-1999; 99WO-US028551.

XX 29-AUG-2001; 2001WO-US027099.

XX 18-JUL-2002; 2002US-00197942.

XX (GENTH) GENENTECH INC.

XX Baker KP, Baton DL, Filvaroff E, Goddard A, Grimaldi JC,

XX Gurney AL, Smith V, Stephan JP, Watanabe CK, Wood WI, Zhang Z,

XX Rong S;

XX WPI; 2003-625487/59.

XX P-PSDB; ADA01232.

XX Novel isolated PRO polypeptides e.g. PRO281 and PRO1560, useful in the

XX preparation of a medicament for treating a condition responsive to PRO

XX polypeptide, and as therapeutic agents e.g. vaccines.

XX Claim 2, Fig 91; 308pp; English.

XX The invention relates to isolated human PRO polypeptides (secreted and

XX transmembrane polypeptides) and the polynucleotides encoding them. The

XX invention also relates to an antibody which specifically binds to a PRO

XX polypeptide, a method for stimulating the release of tumour necrosis

XX factor- α (TNF- α) from human blood, a method for stimulating the

XX proliferation or differentiation of chondrocyte cells and a method for

XX detecting the presence of a tumour in a mammal (e.g. lung, colon, breast,

XX prostate, rectal, kidney and liver tumours). The polynucleotides are

XX useful in molecular biology, including uses as hybridisation probes, in

XX gene therapy. The polynucleotides may also be used in preparing PRO

XX polypeptides by recombinant techniques and in generating either

XX transgenic animals or knock-out animals which are useful in the

XX development and screening of therapeutically useful reagents. The PRO

XX polypeptides or antibodies are used in preparing a medicament for

XX treating a condition responsive to the polypeptides or antibodies, such

XX as tumour, for stimulating and inhibiting proliferation of human

XX microvascular endothelial cells and for inducing endothelial cell tube

XX formation. This sequence represents a human PRO polynucleotide of the

XX invention.

XX Sequence 2597 BP; 688 A; 570 C; 642 G; 697 T; 0 U; 0 Other;

XX Query Match 13.9%; Score 793.8; DB 8; Length 2597;

XX Best Local Similarity 99.7%; Pred. No. 6,1e-181;

XX Mismatches 2; Indels 0; Gaps 0;

XX 579 CGAGTATTTTCCACCATCTCCAGCCGGAAGTACCAAGAACTCTGAGCGGATGGCAT 638

XX 1 CGAGTATTTTCCACCATCTCCAGCCGGAAGTACCAAGAACTCTGAGCGGATGGCAT 60

Db	121	CTTACCACTCATCCCTGACTTGTGATATCTACTATGCTCTATGCTTTTGGACATGGCACTT	180
QY	759	TGTCTACTTTTGGACCCCTCCAACTGAGATGATGTCCTCCACAGGCTTCACCACCMAGA	818
Db	181	TGTCCTACTTTTGGACCCCTCCAACTGAGATGATGTCCTCCACAGGCTTCACCACCMAGA	240
QY	819	GCAGGTATATTCATCCATCCAAAGCTGCTGAGGCTTTGGCAAGAGGACACAGCCTTCAACTCTA	878
Db	241	GCAGGTATATTCATCCATCCAAAGCTGCTGAGGCTTTGGCAAGAGGACACAGCCTTCAACTCTA	300
QY	879	TGTAGAGGTGCCCATTTGCTGTGAGCGCAGTGTGGGTGAGTACCGCCTGCTTGCAGGCTGC	938
Db	301	TGTAGAGGTGCCCATTTGCTGTGAGCGCAGTGTGGGTGAGTACCGCCTGCTTGCAGGCTGC	360
QY	939	CTACTGTATCCAAAGGGGGGGCCGTGCTTGGCAGAGCCCTTGGAGTCCATCCAGATGATGA	998
Db	351	CTACTGTATCCAAAGGGGGGGCCGTGCTTGGCAGAGCCCTTGGAGTCCATCCAGATGATGA	420
QY	999	CCTGCTCTTTCACCGTCTTCTCCAAAGGCCACGAAGCGGAAATGAAATCCCTTGGATGAGTC	1058
Db	421	CCTGCTCTTTCACCGTCTTCTCCAAAGGCCACGAAGCGGAAATGAAATCCCTTGGATGAGTC	480
QY	1059	GGCCCTGTGCATCTTCATCTTGAACAGATTAATACCGCATTAAGAGCGGCTGCAGTGC	1118
Db	481	GGCCCTGTGCATCTTCATCTTGAACAGATTAATACCGCATTAAGAGCGGCTGCAGTGC	540
QY	1119	TTGTTACCGGGGGCGAGGGGACGCTGACCTGTGCTCAAGGTGAAGAGACATCCCTTG	1178
Db	541	TTGTTACCGGGGGCGAGGGGACGCTGACCTGTGCTCAAGGTGAAGAGACATCCCTTG	600
QY	1179	CAGCAGTGGGCTCTTAACCATTTGAGAGATATCTTGTGGGCTGCAGATGAATGCTCCCT	1238
Db	601	CAGCAGTGGGCTCTTAACCATTTGAGAGATATCTTGTGGGCTGCAGATGAATGCTCCCT	660
QY	1239	GGGAGTGTCCGACATGTGTGCTGGAATTCGCTCTTTCACGGAGGACAGGGAACCGCATGAC	1298
Db	661	GGGAGTGTCCGACATGTGTGCTGGAATTCGCTCTTTCACGGAGGACAGGGAACCGCATGAC	720
QY	1299	GTCCTGCATCCGATATGTCTTACAAAGACCATCTGTGGCCTTTGTGTGGGACCAAAAGTGG	1358
Db	721	GTCCTGCATCCGATATGTCTTACAAAGACCATCTGTGGCCTTTGTGTGGGACCAAAAGTGG	780
QY	1359	CAGCTGAAGAAGATCC 1375	
Db	781	CAGCTGAAGAAGGTGC 797	

RESULT	15
ADA43672	
ID	ADA43672 standard; cDNA; 2597 BP.
XX	
XX	ADA43672;
AC	
XX	
DT	20-NOV-2003 (first entry)
XX	
DE	Human CDNA encoding secreted/transmembrane polypeptide PRO34003.

KM ss: gene human; PRO: secreted protein; transmembrane protein;
KM endothelial cell tube formation; chondrocyte cell differentiation;
KM microvascular endothelial cell; tumour; lung tumour; colon tumour;
KM breast tumour; prostate tumour; rectal tumour; kidney tumour;
KM liver tumour; cytostatic; vaccine

OS Homo sapiens.

PN US2003073190-A1.

PD 17-APR-2003.

PF 09-SEP-2002; 2002US-00238283.

PR 01-JUL-1998; 98US-0091358P

02-JUN-1999; 99WO-US012252.
20-JUL-1999; 99US-0144758P.
PR 28-JUL-1999; 99US-0146232P.
PR 25-JUL-1999; 99US-00380137.
PR 30-MAR-2000; 2000WO-US008439.
PR 02-JUN-2000; 2000WO-US015264.
PR 29-AUG-2001; 2001WO-US027099.
PR 18-JUL-2002; 2002US-00197942.
XX
PA (GENTH) GENENTECH INC.
XX
PI Baker KP, Eaton DL, Filvaroff E, Goddard A, Grimaldi JC,
PI Gurney AL, Smith V, Stephan JP, Watanabe CK, Wood WI, Zhang Z,
PI Fong S;
XX WPI: 2003-585304/55.
DR P-PSDB; ADMA3673.
XX
XX New isolated nucleic acid encoding a PRO polypeptide, e.g. PRO20080 or
PT PRO2183, useful in molecular biology, chromosome and gene mapping, in
PT generating antisense RNA and DNA, and in gene therapy.
PS
PS Claim 2; Fig 91; 35app; English.
XX
XX The invention relates to an isolated secreted/transmembrane (PRO)
CC polypeptide, having at least 80% sequence identity to a sequence selected
CC from any one of the 57 amino acid sequences given in specification, or to
CC a sequence encoded by a nucleic acid molecule selected from any one of
CC the nucleic acids deposited under any of the ATCC accession numbers given
CC in specification, or a sequence having at least 80% identity to PRO
CC lacking its associated signal peptide, an extracellular domain of PRO
CC with or without its associated signal peptide. Also included are vectors,
CC transformed host cells, anti-PRO antibodies, the nucleic acids encoding
CC PRO, PRO fusion proteins, inducing endothelial cell tube formation (by
CC administering PRO281, PRO1560, PRO189, PRO499, PRO6308, PRO6000,
CC PRO10275, PRO21207, PRO20933 or PRO34274 polypeptide or its agonist) and
CC an oligonucleotide probe derived from any one of the above nucleotide
CC sequences. PRO6018 polypeptide is useful for stimulating the proliferation
CC and PRO21383 polypeptides are useful for stimulating the proliferation of
CC human microvascular endothelial cells. PRO6077, PRO2487 and PRO6006
CC polypeptides are useful for inhibiting the proliferation of human
CC microvascular endothelial cells. PRO polypeptides are useful for
CC detecting the presence of tumour in a mammal, including tumours of lung,
CC colon, breast, prostate, rectal, kidney and liver. PRO281, PRO1560,
CC PRO189, PRO499, PRO6308, PRO6000, PRO10275, PRO20933 and
CC PRO34274 polypeptides are useful for inducing endothelial cell tube
CC formation. PRO or the antibody are useful in the preparation of a
CC medicament for treating a condition responsive to PRO polypeptide. The
CC oligonucleotide probes are useful for isolating genomic and cDNA
CC nucleotide sequences, for measuring or detecting the expression of an
CC associated gene, and as antisense probes. PRO nucleic acid is useful as a
CC hybridisation probe, in chromosome and gene mapping, in the generation of
CC antisense RNA and DNA, and for the preparation PRO polypeptides. PRO and
CC PRO nucleic acid are useful as therapeutic agents, e.g. vaccines. The
XX present sequence encodes a PRO protein.
XX
XX Sequence 2597 BP; 688 A; 570 C; 642 G; 697 T; 0 U; 0 Other;

Query Match	13.9%	Score 793.8	DB 8	Length 2597
Best Local Similarity	99.7%	Pred. No. 6.1e-181		
Matches 795	0	Mismatches 2	Indels 0	Gaps 0

579 CGAGTATTTTCCACCATCTCCAGCCGGAACCTGACCAAGAACTCTGAGCGGATGGCAT 638

Dh 1 CGAGTATTTCCCAACCATCTCCAGCCGGAACCTGACCAAGACTCTGAGCGGATGCGCAT 60

639 GTTTCGCGTACCGTCTTTCATGATGAGTTCGTGGCCTCGATGATTAAGATCCCTTCGGACAC 698

120

[illegible]

```
Db 121 CTTCAACATCAATCCCTGACTTTGATATCTACTATGTGTGTTTAAAGATGGCAACTT 180
QY 759 TGTCTACTTTTTTGAACCTTCGAACCTGAGATGTGTCTCCACCAAGGCTCCACCAAGGA 818
Db 181 TGTCTACTTTTTTGAACCTTCGAACCTGAGATGTGTCTCCACCAAGGCTCCACCAAGGA 240
QY 819 GAGAGGTATATACATCCAAAGTGTGTGAGGTTTGAAGGAGGACAGAGCTTCAACTCTTA 878
Db 241 GAGAGGTATATACATCCAAAGTGTGTGAGGTTTGAAGGAGGACAGAGCTTCAACTCTTA 300
QY 879 TGTAGAGTGTCCCATTTGAGTGTGAGGCAATGAGGTTGAGTACCGCTGTGTGAGGCTGC 938
Db 301 TGTAGAGTGTCCCATTTGAGTGTGAGGCAATGAGGTTGAGTACCGCTGTGTGAGGCTGC 360
QY 939 CTAACCTGTCCAAAGCGGAGGCGGTGTGTGAGGCAACCTGTGAGTCCATCCAGATGATGA 998
Db 361 CTAACCTGTCCAAAGCGGAGGCGGTGTGTGAGGCAACCTGTGAGTCCATCCAGATGATGA 420
QY 999 CCTGCTCTTCAACCGTCTTCTCCAAAGGAGGCAAGGCAAAATGAAATCCCTGATGATG 1058
Db 421 CCTGCTCTTCAACCGTCTTCTCCAAAGGAGGCAAGGCAAAATGAAATCCCTGATGATG 480
QY 1059 GGGCCCTGTGATCTTATCTTGAAGGAGATTAATGACCGCATTTAAGAGCGGCTGCACTC 1118
Db 481 GGGCCCTGTGATCTTATCTTGAAGGAGATTAATGACCGCATTTAAGAGCGGCTGCACTC 540
QY 1119 TTGTACCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1178
Db 541 TTGTACCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 600
QY 1179 CAGCAGTGGCTCTTAACCATTTGATGATGATGATGATGATGATGATGATGATGATG 1238
Db 601 CAGCAGTGGCTCTTAACCATTTGATGATGATGATGATGATGATGATGATGATGATG 660
QY 1239 GGGAGTGTCCGACATGTGTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1298
Db 661 GGGAGTGTCCGACATGTGTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 720
QY 1299 GTCTGTCAATGCAATATGTCTTAAAGAACCACTCTGTGAGGCTTTTGTGAGGCA 1358
Db 721 GTCTGTCAATGCAATATGTCTTAAAGAACCACTCTGTGAGGCTTTTGTGAGGCA 780
QY 1359 CAAAGCTGAAGAGATCC 1375
Db 781 CAAAGCTGAAGAGATGTC 797
```

Search completed: February 19, 2004, 16:52:25
Job time : 1359 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: February 19, 2004, 16:18:27 : Search time 259 Seconds
(without alignments)
12193.921 Million cell updates/sec

Title: US-09-964-956-12

Perfect score: 5691
Sequence: 1 atgaagcagccctgaa.....gcttagacagctgaataaa 5691

Scoring table: IDENTITY NUC
Gapop 10.0, Gapext 1.0

Searched: 682709 seqs, 277475446 residues

Total number of hits satisfying chosen parameters: 1365418

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued Patents NA:*

1: /cgn2_6/ptodata/2/ina/5A.COMB.seq:*
2: /cgn2_6/ptodata/2/ina/5B.COMB.seq:*
3: /cgn2_6/ptodata/2/ina/6A.COMB.seq:*
4: /cgn2_6/ptodata/2/ina/6B.COMB.seq:*
5: /cgn2_6/ptodata/2/ina/PCTUS.COMB.seq:*
6: /cgn2_6/ptodata/2/ina/backfile1.seq:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	579.2	10.2	2477	4	US-09-907-794A-169
2	579.2	10.2	2477	4	US-09-905-125A-169
3	579.2	10.2	2477	4	US-09-902-775A-169
4	284.2	5.0	3458	4	US-09-023-655-603
5	201.2	3.5	4707	3	US-09-181-706-1
6	201.2	3.5	4707	3	US-09-458-791-1
7	201.2	3.5	4707	3	US-09-458-066-1
8	201.2	3.5	4707	4	US-09-458-065-1
9	86.4	1.5	1024	4	US-09-328-475C-37
10	57.8	1.0	288	4	US-09-313-284A-6281
11	49.6	0.9	7218	1	US-08-233-463-14
12	49.4	0.9	1867	2	US-08-607-509-3
13	49.4	0.9	1867	2	US-08-634-642-3
14	49.4	0.9	1867	2	US-08-989-370-3
15	49.4	0.9	1867	4	US-09-398-169-3
16	49.2	0.9	4626	1	US-08-306-651B-22
17	49.2	0.9	4626	5	PCT-US93-06251-27
18	48.6	0.9	3012	4	US-09-551-974A-94
19	48.6	0.9	3012	4	US-09-556-501A-94
20	48.6	0.9	3012	4	US-09-639-206A-94
21	48.6	0.9	3012	4	US-09-874-923-94
22	48.6	0.9	4917	4	US-09-551-974A-100
23	48.6	0.9	4917	4	US-09-556-501A-100
24	48.6	0.9	4917	4	US-09-639-206A-100
25	48.6	0.9	4917	4	US-09-874-923-100
26	48.6	0.9	4929	4	US-09-551-974A-98
27	48.6	0.9	4929	4	US-09-556-501A-98

28	48.6	0.9	4929	4	US-09-639-206A-98	Sequence 98, Appl
29	48.6	0.9	4929	4	US-09-874-923-98	Sequence 98, Appl
30	48.2	0.8	1241	1	US-08-471-033-39	Sequence 39, Appl
31	48.2	0.8	1241	1	US-08-471-033-42	Sequence 42, Appl
32	48.2	0.8	1241	2	US-08-471-044-35	Sequence 39, Appl
33	48.2	0.8	1241	2	US-08-471-044-42	Sequence 42, Appl
34	48.2	0.8	1241	2	US-08-463-483A-39	Sequence 39, Appl
35	48.2	0.8	1241	2	US-08-463-483A-42	Sequence 42, Appl
36	48.2	0.8	1241	2	US-08-471-046A-39	Sequence 39, Appl
37	48.2	0.8	1241	2	US-08-471-046A-42	Sequence 42, Appl
38	48.2	0.8	1241	2	US-08-470-566B-39	Sequence 39, Appl
39	48.2	0.8	1241	2	US-08-470-566B-42	Sequence 42, Appl
40	48.2	0.8	1241	2	US-08-469-334-39	Sequence 39, Appl
41	48.2	0.8	1241	2	US-08-469-334-42	Sequence 42, Appl
42	48.2	0.8	1241	3	US-09-300-529-39	Sequence 39, Appl
43	48.2	0.8	1241	3	US-09-300-529-42	Sequence 42, Appl
44	48.2	0.8	1358	1	US-08-471-033-45	Sequence 45, Appl
45	48.2	0.8	1358	2	US-08-471-044-45	Sequence 45, Appl

ALIGNMENTS

RESULT 1
US-09-907-794A-169
Sequence 169, Application US/09907794A
Patent No. 6635468
GENERAL INFORMATION:
APPLICANT: Genentech, Inc.
APPLICANT: Ashkenazi, Avi
APPLICANT: Botstein, David
APPLICANT: Desnoyers, Luc
APPLICANT: Eaton, Dan L.
APPLICANT: Ferrara, Napoleone
APPLICANT: Filvaroft, Ellen
APPLICANT: Fong, Sherman
APPLICANT: Gao, Wei-Qiang
APPLICANT: Gerber, Hanspeter
APPLICANT: Gertsen, Mary E.
APPLICANT: Goddard, A.
APPLICANT: Godowski, Paul J.
APPLICANT: Grimaldi, Christopher J.
APPLICANT: Gueney, Austin L.
APPLICANT: Hillan, Kenneth, J.
APPLICANT: Kljavin, Ivar J.
APPLICANT: Mather, Jennie P.
APPLICANT: Pan, James
APPLICANT: Paoni, Nicholas F.
APPLICANT: Roy, Margaret Ann
APPLICANT: Stewart, Timothy A.
APPLICANT: Tumas, Daniel
APPLICANT: Williams, P. Mickey
APPLICANT: Wood, William, I.
TITLE OF INVENTION: Secreted and Transmembrane Polypeptides and Nucleic
FILE REFERENCE: 10466-14
CURRENT APPLICATION NUMBER: US/09/907,794A
CURRENT FILING DATE: 2001-07-17
PRIOR APPLICATION NUMBER: PCT/US00/04414
PRIOR FILING DATE: 2000-02-22
PRIOR APPLICATION NUMBER: US 60/143,048
PRIOR FILING DATE: 1999-07-07
PRIOR APPLICATION NUMBER: US 60/145,698
PRIOR FILING DATE: 1999-07-26
PRIOR APPLICATION NUMBER: US 60/146,222
PRIOR FILING DATE: 1999-07-28
PRIOR APPLICATION NUMBER: PCT/US99/20594
PRIOR FILING DATE: 1999-09-08
PRIOR APPLICATION NUMBER: PCT/US99/20944
PRIOR FILING DATE: 1999-09-13
PRIOR APPLICATION NUMBER: PCT/US99/21090
PRIOR FILING DATE: 1999-09-15
PRIOR APPLICATION NUMBER: PCT/US99/21547

```

; PRIOR FILING DATE: 1999-09-15
; PRIOR APPLICATION NUMBER: PCT/US99/23089
; PRIOR FILING DATE: 1999-10-05
; PRIOR APPLICATION NUMBER: PCT/US99/28214
; PRIOR FILING DATE: 1999-11-29
; PRIOR APPLICATION NUMBER: PCT/US99/28313
; PRIOR FILING DATE: 1999-11-30
; PRIOR APPLICATION NUMBER: PCT/US99/28564
; PRIOR FILING DATE: 1999-12-02
; PRIOR APPLICATION NUMBER: PCT/US99/28565
; PRIOR FILING DATE: 1999-12-02
; PRIOR APPLICATION NUMBER: PCT/US99/30095
; PRIOR FILING DATE: 1999-12-16
; PRIOR APPLICATION NUMBER: PCT/US99/30911
; PRIOR FILING DATE: 1999-12-20
; PRIOR APPLICATION NUMBER: PCT/US99/30999
; PRIOR FILING DATE: 1999-12-20
; PRIOR APPLICATION NUMBER: PCT/US00/00219
; PRIOR FILING DATE: 2000-01-05
; NUMBER OF SEQ ID NOS: 423
; SEQ ID NO 169
; LENGTH: 2477
; TYPE: DNA
; ORGANISM: Homo sapiens
; US-09-907-794A-169

Query Match      10.2%; Score 579.2; DB 4; Length 2477;
Best Local Similarity 67.1%; Pred. No. 1,4e-131;
Matches 837; Conservative 0; Mismatches 408; Indels 3; Gaps 1;

QY 151 TTCAATCACTGCTGCTGATGAGAGAGACACATTACTTTGGGGCCCGTCATCCG 210
DB 976 TTCAACCACTTGACCGCTCCACCAAGGACGCGGCGCTCTATGTGGGGCCCATCAACCG 1035
QY 211 ATTCAAGAGCTCTCCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 270
DB 1036 GTCTATAGAGCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1095
QY 271 GACAAACCCCAAGTTTACCCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 330
DB 1096 GACAAACAGTCTGCTTACCCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1155
QY 331 ACCAAGATGTCACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 390
DB 1156 ACCAAGATGTCACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1215
QY 391 GAGAGAGCTGTCACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 450
DB 1216 GAGAGAGCTGTCACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1275
QY 451 GAGAGCTATATAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 510
DB 1276 GAGAGCTATATAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1335
QY 511 GAGAGCTATATAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 570
DB 1336 GAGAGCTATATAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1395
QY 571 GAGAGAGCTGTCACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 630
DB 1396 GAGAGAGCTGTCACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1455
QY 631 GATGAGAGCTGTCACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 690
DB 1456 TCAGAGAGCTGTCACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1515
QY 691 TCAGAGAGCTGTCACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 750
DB 1516 TCAGAGAGCTGTCACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1575
QY 751 GAGCACTTGTCTACTTTTGAACCTTCAACCTGAGA---TGATGCTTCAACAGAGCTCC 807
DB 1576 GGGGGCTTGTCTACTTTTGAACCTTCAACCTGAGA---TGATGCTTCAACAGAGCTCC 1635

```

```

QY 808 ACCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 867
DB 1636 TCAGAGAGCTTGTCTACTTTTGAACCTTCAACCTGAGA---TGATGCTTCAACAGAGCTCC 1695
QY 868 TTCAACTCTATATAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 927
DB 1696 TTCAACTCTATATAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1755
QY 928 CTGAGAGCTGCTTCAACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 987
DB 1756 CTGAGAGCTGCTTCAACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1815
QY 988 CCAGATGATGACCTGCTTCAACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1047
DB 1816 AGCCAGAGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1875
QY 1048 CTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1107
DB 1876 CCAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1935
QY 1108 CGAGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1167
DB 1936 CGAGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1995
QY 1168 GACATCCCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1227
DB 1996 GACATCCCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2055
QY 1228 AATGCTCCCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1287
DB 2056 AATGCTCCCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2115
QY 1288 GACGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1347
DB 2116 GACGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2175
QY 1348 ACCAAGATGTCACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1395
DB 2176 ACTAAGATGTCACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2223

RESULT 2
US-09-905-125A-169
; Sequence 169, Application US/0905125A
; Patent No. 6664376
; GENERAL INFORMATION:
; APPLICANT: Genentech, Inc.
; APPLICANT: Ashkenazi, Avi
; APPLICANT: Botstein, David
; APPLICANT: Desnoyers, Luc
; APPLICANT: Eaton, Dan L.
; APPLICANT: Ferrara, Napoleone
; APPLICANT: Filvaroff, Ellen
; APPLICANT: Fong, Sherman
; APPLICANT: Gao, Wei-Qiang
; APPLICANT: Gerber, Hanspeter
; APPLICANT: Gertsen, Mary E.
; APPLICANT: Goddard, A.
; APPLICANT: Godowski, Paul J.
; APPLICANT: Grimaldi, Christopher J.
; APPLICANT: Gurney, Austin L.
; APPLICANT: Hillan, Kenneth, J.
; APPLICANT: Kijavlin, Ivar J.
; APPLICANT: Mather, Jennie P.
; APPLICANT: Pan, James
; APPLICANT: Paoni, Nicholas F.
; APPLICANT: Roy, Margaret Ann
; APPLICANT: Stewart, Timothy A.
; APPLICANT: Tumas, Daniel
; APPLICANT: Williams, P. Mickey
; APPLICANT: Wood, William, I.
; TITLE OF INVENTION: Secreted and Transmembrane Polypeptides and Nucleic

```

TITLE OF INVENTION: Acids Encoding the Same
FILE REFERENCE: 10466-14
CURRENT APPLICATION NUMBER: US/09/905,125A
CURRENT FILING DATE: 2001-07-12
PRIOR APPLICATION NUMBER: PCT/US00/04414
PRIOR FILING DATE: 2000-02-22
PRIOR APPLICATION NUMBER: US 60/143,048
PRIOR FILING DATE: 1999-07-07
PRIOR APPLICATION NUMBER: US 60/145,698
PRIOR FILING DATE: 1999-07-26
PRIOR APPLICATION NUMBER: US 60/146,222
PRIOR FILING DATE: 1999-07-28
PRIOR APPLICATION NUMBER: PCT/US99/20594
PRIOR FILING DATE: 1999-09-08
PRIOR APPLICATION NUMBER: PCT/US99/20944
PRIOR FILING DATE: 1999-09-13
PRIOR APPLICATION NUMBER: PCT/US99/21090
PRIOR FILING DATE: 1999-09-15
PRIOR APPLICATION NUMBER: PCT/US99/21547
PRIOR FILING DATE: 1999-09-15
PRIOR APPLICATION NUMBER: PCT/US99/23089
PRIOR FILING DATE: 1999-10-05
PRIOR APPLICATION NUMBER: PCT/US99/28214
PRIOR FILING DATE: 1999-11-29
PRIOR APPLICATION NUMBER: PCT/US99/28313
PRIOR FILING DATE: 1999-11-30
PRIOR APPLICATION NUMBER: PCT/US99/28564
PRIOR FILING DATE: 1999-12-02
PRIOR APPLICATION NUMBER: PCT/US99/28565
PRIOR FILING DATE: 1999-12-02
PRIOR APPLICATION NUMBER: PCT/US99/30095
PRIOR FILING DATE: 1999-12-16
PRIOR APPLICATION NUMBER: PCT/US99/30911
PRIOR FILING DATE: 1999-12-20
PRIOR APPLICATION NUMBER: PCT/US99/30999
PRIOR FILING DATE: 1999-12-20
PRIOR APPLICATION NUMBER: PCT/US00/00219
PRIOR FILING DATE: 2000-01-05
NUMBER OF SEQ ID NOS: 423
SEQ ID NO 169
LENGTH: 2477
TYPE: DNA
ORGANISM: Homo sapiens
US-09-905-125A-169

Query Match 10.2%; Score 579.2; DB 4; Length 2477;
Best Local Similarity 67.1%; Pred. No. 1,4e-131;
Matches 837; Conservative 0; Mismatches 408; Indels 3; Gaps 1;

151 TTCAATACCTGCTGATGATGAGAGACACATTTCTGGGGCCCTCAATCG 210
976 TTCAACACCTTGACCGTCCACAGGAGCGGGCCGTATGATGGGGCCATCAACCG 1035
211 ATTACAGCTCTCCAGCAGCTGAAAGCTTTGATGACGATGAGAGAGGCGGAG 270
1036 GTCTATAGCTGACAGGCAACCTGACCATCAGTGTCTATAGAGAGGGCCAGAG 1095
271 GACAACCCCAAGTTTACCAACCCCGCATGCTCAGATGAGCCCTGACACCC 330
1096 GACAACAAGTCTTACCGCCCTCATCTGTGAGCCCTGACAGAGTCTCAACCTC 1155
331 ACCAACAATGTGAAGATGCTCTCATAGCTCAAGAGAAACAGGCTGATGCTCT 390
1156 ACCAACAATGTGAAGATGCTCTCATAGCTCAAGAGAAACAGGCTGATGCTCT 1215
391 GAGAGCTGTACCAAGCATCTGCAAGCTGAGGCTGAGAGACCTCTTCAAGCTGG 450
1216 GAGAGCTGTACCAAGCATCTGCAAGCTGAGGCTGAGAGACCTCTTCAAGCTGG 1275
451 GAGCTTATATATAGAGAGAGCATATCTGTCAAGTGTCAACGAGAGGCTGAGCTTT 510
1276 GAGCTTATATATAGAGAGAGCATATCTGTCAAGTGTCAACGAGAGGCTGAGCTTT 1335

511 GAGCTGATGCTCTTCTCAACGAACTGGATGACAGCTGTTCATTGCGACGGAGTGGAT 570
1336 GGGGTGATGCTGCTCGAGAGGAGTGGAGATGGCAAGCTCTTACCGGACGGCTGTGGAT 1395
571 GGGAGCCCGAGATATTTTCCACCATCTCCAGCCGGAATAGAACCAAGATCTGAGGCG 630
1396 GGGAGAGAGATTAATCTTCCGACCTCTTCCAGCCGGAAGCTGCGCCGAGACCTGATGCC 1455
631 GATGAGATGCTGCGGATGCTCTTCCATGATGATGATGCTGCGCTGATGATTAATGCTCT 690
1456 TCAGCATGCTGCGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1515
691 TCGACACCTTACCAATCATCTTGAATTTGATATCTATGCTATGCTTTTGAAGCT 750
1516 TCAGACACCTTGGGCGCTGCTTCCACTTTCATCTTTCATCTTTCATCTTTCATCTTTCATCT 1575
751 GGCACCTTGTCTACTTTTGAACCTTCAACCTGAGA---TGTGTCTTCCACGAGCTCC 807
1576 GGGGCTTTGTCTACTTTTCACTTTCATCTTTCATCTTTCATCTTTCATCTTTCATCTTTCATCT 1635
808 ACCAACAAGAGAGAGTGTATATATCAATCAAGCTGCTGAGCTTTGCAAGAGAGACAGCC 867
1636 TCCGCTGAGACCTTCTTCACTTTCATCTTTCATCTTTCATCTTTCATCTTTCATCTTTCATCT 1695
868 TTCAACTTCTTGTATAGAGTGGCCCATTTGGCTGTGAGGCGAGTGGAGTGAATGAGCTGCT 927
1696 TTCACTTATATGCTGTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 1755
928 CTGACAGCTGCTTCACTGTCCAAAGCGGGCGGCTTGGAGAGACCTTGAAGTCCAT 987
1756 CTGACAGCTGCTTCACTGTCCAAAGCGGGCGGCTTGGAGAGACCTTGAAGTCCAT 1815
988 CCAATATGACCTGCTTCTTCACTGTCCAAAGCGGGCGGCTTGGAGAGACCTTGAAGTCCAT 1047
1816 AGCCAGAGAGATGATCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 1875
1048 CTGATGATGCTGCGCTTGTGATCTTCACTTTCATCTTTCATCTTTCATCTTTCATCTTTCATCT 1107
1876 CCGATGATCTGCGCTTGTGATCTTCACTTTCATCTTTCATCTTTCATCTTTCATCTTTCATCT 1935
1108 CCGCTGATCTTGTATACCGGGCGAGGGACCTTGAACCTGAGCTGCTTGAAGTGAAG 1167
1936 CCGCTGATCTTGTATACCGGGCGAGGGACCTTGAACCTGAGCTGCTTGAAGTGAAG 1995
1168 GACATCCCTGACAGAGTGCCTTCACTTTCATCTTTCATCTTTCATCTTTCATCTTTCATCTTTCATCT 1227
1996 GACATCCCTGACAGAGTGCCTTCACTTTCATCTTTCATCTTTCATCTTTCATCTTTCATCTTTCATCT 2055
1228 AATGCTCCCTGAGAGTGTCCGATATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1287
2056 AACCAAGCCCTGAGAGTGTCCGATATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2115
1288 GACGCGAT 2147
2116 GACGCGAT 2175
1348 ACCAAAGTGCAGAGCTGAGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1395
2176 ACTAAGATGCGCAAGCTGAGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2223

RESULT 3

US-09-902-775A-169
Sequence 169, Application US/09902775A
Patent No. 6686451
GENERAL INFORMATION:
APPLICANT: Genentech, Inc.
APPLICANT: Askenazi, Avi
APPLICANT: Botstein, David
APPLICANT: Desrochers, Luc
APPLICANT: Eaton, Dan L.
APPLICANT: Ferrara, Napoleone
APPLICANT: Filvaroff, Ellen

```

APPLICANT: Fong, Sherman
APPLICANT: Gao, Wei-Qiang
APPLICANT: Gerber, Hanspeter
APPLICANT: Gerritsen, Mary E.
APPLICANT: Goddard, A.
APPLICANT: Godowski, Paul J.
APPLICANT: Grimaldi, Christopher J.
APPLICANT: Gurney, Austin L.
APPLICANT: Hillan, Kenneth, J.
APPLICANT: Kijaviri, Ivar J.
APPLICANT: Mather, Jennie P.
APPLICANT: Pan, James
APPLICANT: Paoni, Nicholas F.
APPLICANT: Roy, Margaret Ann
APPLICANT: Stewart, Timothy A.
APPLICANT: Tumas, Daniel
APPLICANT: Williams, P. Mickey
APPLICANT: Wood, William, I.
TITLE OF INVENTION: Secreted and Transmembrane Polypeptides and Nucleic
TITLE OF INVENTION: Acids Encoding the Same
FILE REFERENCE: 10466-14
CURRENT APPLICATION NUMBER: US/09/902,775A
PRIOR FILING DATE: 2001-07-10
PRIOR APPLICATION NUMBER: PCT/US00/04414
PRIOR FILING DATE: 2000-02-22
PRIOR APPLICATION NUMBER: US 60/143,048
PRIOR FILING DATE: 1999-07-07
PRIOR APPLICATION NUMBER: US 60/145,698
PRIOR FILING DATE: 1999-07-26
PRIOR APPLICATION NUMBER: US 60/146,222
PRIOR FILING DATE: 1999-07-28
PRIOR APPLICATION NUMBER: PCT/US99/20594
PRIOR FILING DATE: 1999-09-08
PRIOR APPLICATION NUMBER: PCT/US99/20944
PRIOR FILING DATE: 1999-09-13
PRIOR APPLICATION NUMBER: PCT/US99/21090
PRIOR FILING DATE: 1999-09-15
PRIOR APPLICATION NUMBER: PCT/US99/21547
PRIOR FILING DATE: 1999-09-15
PRIOR APPLICATION NUMBER: PCT/US99/23089
PRIOR FILING DATE: 1999-10-05
PRIOR APPLICATION NUMBER: PCT/US99/28214
PRIOR FILING DATE: 1999-11-29
PRIOR APPLICATION NUMBER: PCT/US99/28313
PRIOR FILING DATE: 1999-11-30
PRIOR APPLICATION NUMBER: PCT/US99/28564
PRIOR FILING DATE: 1999-12-02
PRIOR APPLICATION NUMBER: PCT/US99/28565
PRIOR FILING DATE: 1999-12-02
PRIOR APPLICATION NUMBER: PCT/US99/30095
PRIOR FILING DATE: 1999-12-16
PRIOR APPLICATION NUMBER: PCT/US99/30911
PRIOR FILING DATE: 1999-12-20
PRIOR APPLICATION NUMBER: PCT/US99/30999
PRIOR FILING DATE: 1999-12-20
PRIOR APPLICATION NUMBER: PCT/US00/00219
PRIOR FILING DATE: 2000-01-05
NUMBER OF SEQ ID NOS: 423
SEQ ID NO 169
LENGTH: 2477
TYPE: DNA
ORGANISM: Homo sapiens
US-09-902-775A-169

```

Query Match	Similarity	Score	579.2	DB 4	Length	2477
Best Local	Similarity	67.1%	Pred. No.	1.4e-131		
Matches	837	Conservative	0	Mismatches	408	Indels
						Gaps
						1
OY	151	TTCAATCACTCGTGTGTGATGAGAGACAGACACATTACTTGGGGGCGCGTAATTCG				210
Db	976	TTCAACCACTTGACCGTCCACCCAGAGACGGGGCGCGTCTTGTGTGGGGGCATACACCG				1035
OY	211	ATTTCACAGACTCTCCAGGAGACCTCTGAAGTCTTGTGTGACGCATGAGACAGGGCCGGAAG				270

Db	1036	GTCTATTAGCTGACAGGCAACCTGACCATCAGGTGGCTCTATAAGACAGGGGCCAAGAG	1095
Qy	271	GACAAACCCCAAGTTTACCCACCCCGCATGTCGACACTGCAATGAGCCCTGACCAAC	330
Db	1096	GACAAACAAGTCTGGTTACCCGCCCTCATGTGCGAGCCCTGCAGCGAAGTGTACACCTC	1155
Qy	331	ACCAACAATGTGCAACAAGATGCTCCTCATAGACATCAAGAGAAACAGAGCTGATTCGCTGT	390
Db	1156	ACCAACAATGTGCAACAAGCTGTCTCATATGATGATCTAGAAACCGCTGTGGCTGT	1215
Qy	391	GGAGCCTGTACCAAGGCATCTGCAAGCTGTGAGGCTGAGAGACCTTTCAGAGTGGG	450
Db	1216	GGAGCCTGTACCAAGGGGGTGTGCAACCTGTGTGGGTGATGACCTTTCATCTGTGTG	1275
Qy	451	GAGCCTATATCAATGAAGAGAGCATATCTGTCAAGGTGTCAACGAGACGGCTCAGTCTT	510
Db	1276	GAGCCTATCCCAAGAAAGAGACATTAAGCTGTCAACAGAGAGGGGACCATGTAC	1335
Qy	511	GAGTGTATCGTCTCTACAGCAACCTGATGACAAAGTGTTCATTGCCACGGCAGTGGAT	570
Db	1336	GGGTGTATGTGGCCTCTGAGGGATGAGGATGCAAGCTCTTCATCGGACCGCTGTGGAT	1395
Qy	571	GGGAAGCCCGAGTATTTTCCACCATCTCCAGCCGGAAATGATCAAGAACTCTGAGCG	630
Db	1396	GGGAAGCGAGATTAATCTCCCACTCTGTCCAGCCGGAAAGTGTCCCGAGACCTGTAGTCC	1455
Qy	631	GATGCATGTTCCGGTAGCTTCCATGATGATGTTCTGTGGCTTGATATTAAGTCCCT	690
Db	1456	TCAGCATAGCTCGACTATGAGCTACACAGCCATTGTGTCTCTCATCAAGATCCCT	1515
Qy	691	TCGACACCTTACCATCATCCCGAGCTTTGATATCTACTATGTCTATAGTTTAGCAGT	750
Db	1516	TCAGACCTCTGGCCCTGTGTCTCCACCTTGACATCTTCAACTATAGAGCTTGTGATGT	1575
Qy	751	GACACATTGTCTACTTTTGAACCTTCAACCTGAGA--TGTGTCTCCACAGGCTTC	807
Db	1576	GGGGGCTTGTCTCTTCTTCTCATGCTCCAGCCCGAGACCCCTGAGGGGTGACATCAAC	1635
Qy	808	ACCCACAGAGAGAGGTGTATACATCCAAAGCTCGAGGCTTTGCAAGAGACACAGCC	867
Db	1636	TCGGGTGAGACCTTTCTACCTCACCTACAGCATGTGTGGGCTCTGCAAGATGACCCCAAG	1699
Qy	868	TTCAACTCTTATGTAGAGGTGCCATTGGCTGTGAGCGCAGTGGGGTGTGAGTACCGCTG	927
Db	1696	TTCCACTCATAGTGTCCCTGCGCTTCGGCTGTGACCCCGGGCCGGGGTGAATAACGGCTTC	1755
Qy	928	CTGCAAGCTGCTTACTCTGTCCAAAGCGGGGGCGGTGTGGCAGAGCCCTTGAATGCAT	987
Db	1756	CTGCAAGCTGCTTACTCTGTCCAAAGCTGTGGAGTACTGTGGCCAGGCTTCAATATCAC	1811
Qy	988	CCAGATGATGACCTGCTCTTCAACGCTCTTTCCAAGGCGCCAGAGCCGAAATAGAATTC	104
Db	1816	AGCCAGGACGATGTACTCTTGGCCATCTTCTCCAAAGGCGAGAAACAGATATCACCAACCG	1877
Qy	1048	CTGATGTAGTGGCCCTGTGATCTTCAATCTTGAAGACAGATPAAATGACCGCATTTAAGAG	1107
Db	1876	CCCGATGATCTGCGCTGTGTGCTTCCATCTTCGGGCAATCACTTGAGATCAAGAG	1933
Qy	1108	CGGCTGCACTTTTGTATCCGGGGCAGAGGACGCTGACCTTGCGCTGTCAAGTGAAG	116
Db	1936	CGCTGCAATCTGTCTACACAGGGCAGGGCAACTGAGGCTCAACTGTGCTGTGGGAAG	1999
Qy	1168	GACATCTCCCTGAGAGAGTGGCTTAAACATATGACATTAATCTTGTGTGGCTGTGACATG	122
Db	1996	GAGCTCACTGTACCGAAGGCGCTGTCCCAATCATATTAATCTTGTGTGAGCTGACATC	205
Qy	1228	AATGCTCCCTGGAGTGTCCGACATGTGTGTGAATATCCGTCTTACGGAGGACAG	128
Db	2056	AACCAAGCTCTGGAGGCTCAACTCCAGTGAAGGGCTGTGACCTGTATCACACACAGAG	211
Qy	1288	GACCGATACGTCTGTCAATGCATATGTCTTCAAGAAACACATCTCTGGCTTTGTGGCG	134

Db 2116 GACCGCATGACCTCTGTGGCTCTCTACGTTTAAACGGCTACAGGCTGTTTGTGGGG 2175
 QY 1348 ACCAAGAGTGGCAGTGAAGAAATCCGGGTGATGAGCCAGGGGC 1395
 Db 2176 ACTAAGAGTGGCAGTGAAGAAATGTTAGAGTCTATGATTCAGATGC 2223

RESULT 4

US-09-023-655-603
 ; Sequence 603, Application US/09023655
 ; Patent No. 6607879

GENERAL INFORMATION:

APPLICANT: Cocks, Benjamin G.
 APPLICANT: Susan G. Stuart
 APPLICANT: Jeffrey J. Seilhamer
 TITLE OF INVENTION: COMPOSITION FOR THE DETECTION OF BLOOD CELL GENE
 NUMBER OF SEQUENCES: 1508
 CORRESPONDENCE ADDRESS:
 ADDRESSEE: INCYTE PHARMACEUTICALS, INC.
 STREET: 3174 PORTER DRIVE
 CITY: PALO ALTO
 STATE: CALIFORNIA
 COUNTRY: USA
 ZIP: 94304

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk
 COMPUTER: IBM PC compatible
 OPERATING SYSTEM: PC-DOS/MS-DOS
 SOFTWARE: Word Perfect 6.1 for Windows/MS-DOS 6.2
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/09/023,655
 FILING DATE: HEREMITH

CLASSIFICATION:

PRIOR APPLICATION DATA:
 APPLICATION NUMBER:

FILING DATE:

CLASSIFICATION:
 ATTORNEY/AGENT INFORMATION:

NAME: Zeller, Karen J.
 REGISTRATION NUMBER: 37,071
 REFERENCE/DOCKET NUMBER: PA-0001 US
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: (650) 855-0555
 TELEFAX: (650) 845-4166
 INFORMATION FOR SEQ ID NO: 603:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 3458 base pairs
 TYPE: nucleic acid
 STRANDEDNESS: single
 TOPOLOGY: linear
 IMMEDIATE SOURCE:
 LIBRARY: THYMON04
 CLONE: 3191066

US-09-023-655-603

Query Match 5.0%; Score 284.2; DB 4; Length 3458;
 Best Local Similarity 64.4%; Pred. No. 2e-59;
 Matches 440; Conservative 0; Mismatches 237; Indels 6; Gaps 1;

QY 5001 CCGGGGACAGATGCTGTCTAATCTACCTGACCGGACCTCGGACCACTAAGGGAC 5060
 Db 1658 CATTCCAGAGAGTCTCCCGAATCTACCTGACCGGACCTGCTCTCCACCAAGGGAC 1717
 QY 5061 ACTGAGAGATTGTGATGACCTCTTTGAGACCACTTTACAGACGGCAACCGTGGCTC 5120
 Db 1718 GTTGAGAGATTCTGATGACCTGTTCAGAGCCATTCTGAGTA-----TCCGTGAAGA 1771
 QY 5121 TGGCCCTGGCCGATCATAGATGATGTTTACCTCCGTGATGAGGAGCGATGATAACA 5180
 Db 1772 CAAGCCCCCACTGGCTGCAANNACTTTTCGACTTCCGTGAGANGAGCGCTGAGAGAG 1831
 QY 5181 TGGCATTGATGACCGCAGCTCCGCCATACCTGAGAGAGCAATTGCTTGGCCCTGAGGTT 5240

Db 1832 GGGAAATCTCCGACCCCGACACCTTACACATCTGTGAAGACCAAGCTTCTCCGGTT 1891
 QY 5241 TTGGGTACATGATGATGAGAACCCGAGTTTGTGTTGACATCTCATAGAACGATCAC 5300
 Db 1892 CTGGGTGACATCTCTGATGAGAACCCCGAGTTTGTGTTGACATCTCATAGAACGATCAC 1951
 QY 5301 AGACGCTGCTCTCTGTGTGCTGATGACCTTCAATGACCTTGTCTTCCAGTCAAGCA 5360
 Db 1952 CGACGCTGCTCTTCAATGATCTCCGACAGGCTTTCATGACGCTGCTCATCTCACT 2011
 QY 5361 CCGGCTGGGCAAGACCTCGGCTTCCAAAGAGTGTGTATGACCAAGACATCCCGAGCT 5420
 Db 2012 GCACTGGGCAAGATTTGCAACCAAGAGTCTCTTCAAGAGATTTCTTGA 2071
 QY 5421 CAAGAAATGGGTGAGAGATTAATCTCAAGATAGGAAAGATCCAGCATCGACCA 5480
 Db 2072 CCGAAGATGCTGACAGGCTACTAACAAGCAGATCCAGGACATAGACCGCTCAGCGAGCA 2131
 QY 5481 AGACATGAAGCATCTGCTGCTGAGAGTCCCGGATGACATGATGATGATTAACCAT 5540
 Db 2132 AGAGATGAATGCCATCTGCGCAGAGGTGAGAGAAATACAGAAATGATTCATCAACCA 2191
 QY 5541 GAGTGCACTCTCAGAGATCTTCTCTATGTGCGCAATACAGGAGAGATCTTGGAC 5600
 Db 2192 TGTGGCATGAGCAGATTTATTAAGTACGCAAGAGGTATCGGCCGACATGAGCGCG 2251
 QY 5601 TCTGGACACGATGACCACTGTGTGGAGAGCAAACTGGCTTAATAACTAGAAATCAT 5660
 Db 2252 GCTGGAGGCGCAACCCCAAGCGCCGAGGAGACAACTGACAGACAAATTTGAGGAGTGT 2311
 QY 5661 AACCTCATGCTTAGACAGCT 5683
 Db 2312 GGTCTTATGAGAGACAACTCT 2334

RESULT 5

US-09-181-706-1
 ; Sequence 1, Application US/09181706
 ; Patent No. 6130068

GENERAL INFORMATION:

APPLICANT: Melanie K. Spriggs, Michael R. Comeau,
 APPLICANT: Robert F. Dubose, Richard S. Johnson
 TITLE OF INVENTION: VIRAL ENCODED SEMAPHORIN PROTEIN
 NUMBER OF SEQUENCES: 10
 CORRESPONDENCE ADDRESS:

ADDRESSEE: Janis C. Henry
 STREET: 51 University St.
 CITY: Seattle
 STATE: WA
 COUNTRY: USA

ZIP: 98101

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk
 COMPUTER: IBM PC compatible
 OPERATING SYSTEM: PC-DOS/MS-DOS
 SOFTWARE: Patemlin Release #1.0, Version #1.30
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/09/181,706
 FILING DATE: October 28, 1998
 CLASSIFICATION:

PRIOR APPLICATION DATA:

APPLICATION NUMBER: US 08/958,598 (converted to a
 APPLICATION NUMBER: Provisional, see below)
 FILING DATE: October 28, 1997

CLASSIFICATION:

PRIOR APPLICATION DATA:
 APPLICATION NUMBER: --to be assigned-- (USN 08/958,598
 APPLICATION NUMBER: conversion to Provisional application)
 FILING DATE: October 26, 1998

CLASSIFICATION:
 ATTORNEY/AGENT INFORMATION:

CORRESPONDENCE ADDRESS:
ADDRESSEE: Janis C. Henry
STREET: 51 University St.
CITY: Seattle
STATE: WA
COUNTRY: US
ZIP: 98101
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: MS-DOS/Windows 95
SOFTWARE: Word for Windows 95, 7.0a
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/458,791
FILING DATE: 10-Dec-1999
CLASSIFICATION: <Unknown>
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/958,598
FILING DATE: 28-Oct-1997
ATTORNEY/AGENT INFORMATION:
NAME: Henry, Janis C
REGISTRATION NUMBER: 34,347
REFERENCE/DOCKET NUMBER: 2631
TELECOMMUNICATION INFORMATION:
TELEPHONE: (206)470-4189
TELEFAX: (206)233-0644
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 4707 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
MOLECULE TYPE: cDNA
HYPOTHETICAL: NO
ANTI-SENSE: NO
FEATURE:
NAME/KEY: CDS
LOCATION: 1..4707
SEQUENCE DESCRIPTION: SEQ ID NO: 1:
US-09-458-791-1
Query Match 3.5%; Score 201.2; DB: 3; Length 4707;
Best Local Similarity 49.1%; Pred. No. 4.3e-39;
Matches 824; Conservative 0; Mismatches 728; Indels 126; Gaps 6;

4355 AGTGTGCGGAGGCCCTCTTCCCTGTTCTGTGCGCATCAGACAGATGAGAGG 4414
4352 AGACTGTGGAGAGGCCCTTCTATTGCTGTGAGAGCTGACCAAGAAATTAAACAAG 3511
4415 GCCCATTTGAGCCATCAGGGCGAGGCCGCTACTCTTTAGAGGAGACAGCTCATCC 4474
3512 GTCCCGTAGATGATATCATCTTGCAAGGCCCTGTACACATTAATGAGACCTGCTGTGT 3571
4475 GCCACAGATTGACTACAAAACCTGTGCTGAGCTGTGACGCCAGACATGCCAACA 4534
3572 GCGAGGTTCCGGAATTGACTGATCTGTGGATTAACGTGCTTTGAAAAAATCCCGGAA 3631
4535 GC-----CCCGAGTCCAGTAAAGATCCTCACTGTGACACATCA 4576
3632 ACGAGGTGAGAGTGTCTGTGGAATATTTCAATGCAATGTTCTGACTGTGACACATTG 3691
4577 CTAGGTCAAGAGAGATTCGTGATGCCATCTTCAAGATGTGCTGCCACCGGC 4636
3692 GCCAAGCCAAAGAAAGTTTCCAGACATCTTAAGCAAAATGGCTCTCTTATGAGAC 3751
4637 CCAAAGCTGCAAGATATGATCTGAGTGGCGACAAAGAGTGGCGAAGATGATCTTGC 4696
3752 TTCACTTAATGAATATGTCTTGAGCTTCAATGGGACACAGCAAGAAAGCTTCGCG 3811
4697 AGGATGAGACATCAACCAAGATTGAGATTTGGAACGACTGACACACTGGGCC 4756
3812 ACATGCAAGTTTCTCCGTGATCTTGAAGATGATCCAGAGTAAACCAATTGGCC 3871
4757 ACTACAGATGCAAGATGTTCCGTGTGTGAGTATGATTCACAGAGTGAACGCTATA 4816
3872 ACTATGAGATATCAATGATGATCTATPAAGTCTTTAAGAGATAGCAATTTACTT 3931
4817 ACGAGTGAACAATCCACCGCTTCCAGGACTCAGCAAGTAATATTAATAACATGATCC 4876
3932 CAGATGTGAGATATCC-----GAGTACCACTGCGCATTTGATTTA----- 3972
4877 GGTACACGGGAGCCCGACAGCTCGCTCAGGACACTATGATATCTCTGACCTGG 4936
3973 ----- 3972
4937 AGATGAGATCAAGATGTGSCACTAGTGAAGAACACAGACAGAGACCAAGAGG 4996
3973 -----CGAGTTTGGAAGATTCACAGATGTGCAAGAAAGAAC-----ATC 4015
4997 GGGACCGGGGGAGCAAGATGTGTGTAATCTACTGACCCGACTCTGGGCCACTAAG 5056
4016 GAGGAGACACAAGATTCAAAGTAAAGAAATGTATGTGACAAAGCTGTGACACAGG 4075
5057 GCAACCTGCAAGAGTTTGTGATGACTCTTTGAGACCATTTGACAGCGACACCGTG 5116
4076 TGGCAATTCATCTGTGCTTAAAAAATTGAGAGCATTTGAGATTACCAACACGCA 4135
5117 GCTGTGCGCTGCGCCGTCATCAAGTACATGTTTGACTTCTGGATGAGAGGCTGATA 5176
4136 GAGCT-----CAATTTCTATAAATCTTTTGACTTTTGGACGCCAGGCTGAAA 4189
5177 AACATGGAATTCATGACCGGACGTCGSCATACCTGTGAAGACGAATTGCTGCCCTGA 5236
4190 ACAAATAATCAAGATCTGACGTCGACATATTGGAATAAACAAGCCTTCTCTTC 4249
5237 GGTTTTGGGTCAATGATCAAGAACCGGAGTTTGTCTTGAATCATATAGAACACACA 5296
4250 GCTTGTGGATTAACCTCTGAAGAACCTCAGTTTGTCTTGAATTAAGAACACACAC 4309
5297 TCACAGAGCGCTGCTCTGTGTGTGCTCAGACCTTACGACTTCTGTCCAGCTGAG 5356
4310 ATATAGAGGCTGTTTGTGAGATTTGCCAGGCAATTCATGATGCAATTTCTCTCAGAG 4369
5357 AGCACCGGCTGGGCAAGACATCGCCCTCCAAACAGCTCTGTATAGCCAGAGCAATCCCA 5416
4370 AGCAGCACTAGGAGAGAGACCAACTAATAGCTTCTATAGCCAAAGATATCCCA 4429
5417 GCTACAGAAATTGGGTGAGAGGTATTACTGACACTAGGAGAGATGCGACGATCAGCG 5476

Db 4430 CCTACCAAGAGAAAGTAAATCTTATTTCACAAAGCAATCAAGGATTTTGCTCCATCTGTCT 4488
 QY 5477 ACCAAGACATGAAGCATACCTGGCTGAGCGTCCGAGTGCATGAAATGATTTCACA 5538
 Db 4490 CCTCAAAATGAGAGATTTTAACTCAGAGATTAAAGAACATGAAATGATTTAAATG 4549
 QY 5537 CCATGAGTGCATCTCAGAGATCTTCTCCTATGTGGCAATACAGCGAGAGATCT 5594
 Db 4550 AAGAGTGGCCTTGACAGAAATTTTCAATACATCTGAAATTTTGTGATGATTTCT 4607

```

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60
61
62
63
64
65
66
67
68
69
70
71
72
73
74
75
76
77
78
79
80
81
82
83
84
85
86
87
88
89
90
91
92
93
94
95
96
97
98
99
100
101
102
103
104
105
106
107
108
109
110
111
112
113
114
115
116
117
118
119
120
121
122
123
124
125
126
127
128
129
130
131
132
133
134
135
136
137
138
139
140
141
142
143
144
145
146
147
148
149
150
151
152
153
154
155
156
157
158
159
160
161
162
163
164
165
166
167
168
169
170
171
172
173
174
175
176
177
178
179
180
181
182
183
184
185
186
187
188
189
190
191
192
193
194
195
196
197
198
199
200
201
202
203
204
205
206
207
208
209
210
211
212
213
214
215
216
217
218
219
220
221
222
223
224
225
226
227
228
229
230
231
232
233
234
235
236
237
238
239
240
241
242
243
244
245
246
247
248
249
250
251
252
253
254
255
256
257
258
259
260
261
262
263
264
265
266
267
268
269
270
271
272
273
274
275
276
277
278
279
280
281
282
283
284
285
286
287
288
289
290
291
292
293
294
295
296
297
298
299
300
301
302
303
304
305
306
307
308
309
310
311
312
313
314
315
316
317
318
319
320
321
322
323
324
325
326
327
328
329
330
331
332
333
334
335
336
337
338
339
340
341
342
343
344
345
346
347
348
349
350
351
352
353
354
355
356
357
358
359
360
361
362
363
364
365
366
367
368
369
370
371
372
373
374
375
376
377
378
379
380
381
382
383
384
385
386
387
388
389
390
391
392
393
394
395
396
397
398
399
400
401
402
403
404
405
406
407
408
409
410
411
412
413
414
415
416
417
418
419
420
421
422
423
424
425
426
427
428
429
430
431
432
433
434
435
436
437
438
439
440
441
442
443
444
445
446
447
448
449
450
451
452
453
454
455
456
457
458
459
460
461
462
463
464
465
466
467
468
469
470
471
472
473
474
475
476
477
478
479
480
481
482
483
484
485
486
487
488
489
490
491
492
493
494
495
496
497
498
499
500
501
502
503
504
505
506
507
508
509
510
511
512
513
514
515
516
517
518
519
520
521
522
523
524
525
526
527
528
529
530
531
532
533
534
535
536
537
538
539
540
541
542
543
544
545
546
547
548
549
550
551
552
553
554
555
556
557
558
559
560
561
562
563
564
565
566
567
568
569
570
571
572
573
574
575
576
577
578
579
580
581
582
583
584
585
586
587
588
589
590
591
592
593
594
595
596
597
598
599
600
601
602
603
604
605
606
607
608
609
610
611
612
613
614
615
616
617
618
619
620
621
622
623
624
625
626
627
628
629
630
631
632
633
634
635
636
637
638
639
640
641
642
643
644
645
646
647
648
649
650
651
652
653
654
655
656
657
658
659
660
661
662
663
664
665
666
667
668
669
670
671
672
673
674
675
676
677
678
679
680
681
682
683
684
685
686
687
688
689
690
691
692
693
694
695
696
697
698
699
700
701
702
703
704
705
706
707
708
709
710
711
712
713
714
715
716
717
718
719
720
721
722
723
724
725
726
727
728
729
730
731
732
733
734
735
736
737
738
739
740
741
742
743
744
745
746
747
748
749
750
751
752
753
754
755
756
757
758
759
760
761
762
763
764
765
766
767
768
769
770
771
772
773
774
775
776
777
778
779
780
781
782
783
784
785
786
787
788
789
790
791
792
793
794
795
796
797
798
799
800
801
802
803
804
805
806
807
808
809
810
811
812
813
814
815
816
817
818
819
820
821
822
823
824
825
826
827
828
829
830
831
832
833
834
835
836
837
838
839
840
841
842
843
844
845
846
847
848
849
850
851
852
853
854
855
856
857
858
859
860
861
862
863
864
865
866
867
868
869
870
871
872
873
874
875
876
877
878
879
880
881
882
883
884
885
886
887
888
889
890
891
892
893
894
895
896
897
898
899
900
901
902
903
904
905
906
907
908
909
910
911
912
913
914
915
916
917
918
919
920
921
922
923
924
925
926
927
928
929
930
931
932
933
934
935
936
937
938
939
940
941
942
943
944
945
946
947
948
949
950
951
952
953
954
955
956
957
958
959
960
961
962
963
964
965
966
967
968
969
970
971
972
973
974
975
976
977
978
979
980
981
982
983
984
985
986
987
988
989
990
991
992
993
994
995
996
997
998
999
1000
1001
1002
1003
1004
1005
1006
1007
1008
1009
1010
1011
1012
1013
1014
1015
1016
1017
1018
1019
1020
1021
1022
1023
1024
1025
1026
1027
1028
1029
1030
1031
1032
1033
1034
1035
1036
1037
1038
1039
1040

```

Query Match	3.5%;	Score 201.2;	DB 3;	Length 4707;
Best Local Similarity	49.1%;	Pred. No. 4.3e-39;		
Matches 824;	Conservative	0;	Mismatches 728;	Indels 126;
				Gaps 6

QY 3935 TTTCGGTCTCGGACCTATAGACCTTACACATCCGGGTCGCTCTTCCCGAGATTAAGAC 399

Db 3038 TTCCCTTCCTGTGATACACACATTTTGTCTTAGAACCTTCTTCCCTGACGACGTGCT 309

QY 3995 ACCCTGTCTCCCGGACCTTTAGAGTCTCCGGGCTACCCGCGAGAGAGCTGTGACAAAGCC 405

Db 3098 TCACCCACATCTTCTACTGAAAGATTCATTAACGAGACGCCACACACAGATGTAAGTCT 315

[illegible]

Db 4136 GAGCT-----CCATTGCTATTAATACTTTTGAAGCCCGAGGCTGAAA 4189
 Qy 5177 AACATGACATTCATGACCCGACGTCGCCATACCTGGAAGACAAATGCTGCCCTGA 5236
 Db 4190 ACAAATAATCAAGATCCGAGCTGTACATTTTGGAAAAAACAACGCTTCCTTC 4249
 Qy 5237 GGTTCGGTCAACATGATCAAGAACCCGAGTTTGTGTTGATCCATTAAGACAGA 5296
 Db 4250 GCTTCGTGTAACATCTGTAAGAACCCCTGAGTTTGTGATTAAGAACACAC 4309
 Qy 5297 TCACAGACGCTGCTCTGTGTGAGCTCAGACCTTACATGACCTTGTCCACGTCAG 5356
 Db 4310 ATATGACGCTGTTTGTCAATGATGCTCCAGGACCTTACATGATGATTTCTCAG 4369
 Qy 5357 AGCACCGCTGGGCAAGACTCGCCCTCCAAACAGCTGTATGCCAAGACATCCCA 5416
 Db 4370 AGCAGCAACTAGGAGGAAGAGCAACAACTAATAGCTTCTATGCAAGATATCCAA 4429
 Qy 5417 GCTACAGATTTGGGTGAGAGGATTAATCAATAGAGGAGGAGATGCCAGCATAGCG 5476
 Db 4430 CTTACAAAGAGAGATTAATCTTATTAACAAGCATCAGGATTTGCTCCATTTGCT 4489
 Qy 5477 ACCAAGCATGAACGATACCTGCTGAGCAGTCCCGGATGACATGAATGAGTTCA 5536
 Db 4490 CCTCAGAAATGGAAGATTTTAACTCAGGATCTAAGAAACATGAATGAATTTAATG 4549
 Qy 5537 CCATGATGCACTCTCAGAGATCTTCTCTATGTGGGCAATACAGGAGAGATGCT 5594
 Db 4550 AAGAGTGCGCTTGACAGAAATTTACAAATACATGTAATTTTGTGATGATTC 4607

RESULT 8

US-09-459-065-1
 Sequence 1, Application US/09459065

Patent No. 6562949
 GENERAL INFORMATION:

APPLICANT: Spitzers, Melanie
 TITLE OF INVENTION: VIRAL ENCODED SEMAPHORIN PROTEIN
 NUMBER OF SEQUENCES: 10
 CORRESPONDENCE ADDRESS: 10
 ADDRESS: 51 University St.
 STREET: 51 University St.
 CITY: Seattle
 STATE: WA
 COUNTRY: US

ZIP: 98101
 COMPUTER READABLE FORM:
 MEDIUM TYPE: Floppy disk
 COMPUTER: IBM PC compatible
 OPERATING SYSTEM: MS-DOS/Windows 95
 SOFTWARE: Word for Windows 95, 7.0a

CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/09/459,065
 FILING DATE:
 CLASSIFICATION:
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: 08/958,598

FILING DATE:
 ATTORNEY/AGENT INFORMATION:
 NAME: Henry, Janis C
 REGISTRATION NUMBER: 34,347
 REFERENCE/DOCKET NUMBER: 2631
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: (206)470-4189
 TELEFAX: (206)233-0644

INFORMATION FOR SEQ ID NO: 1:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 4707 base pairs
 TYPE: nucleic acid
 STRANDEDNESS: double
 TOPOLOGY: linear

MOLECULE TYPE: CDNA
 HYPOTHETICAL: NO
 ANTI-SENSE: NO
 FEATURE:
 NAME/KEY: CDS
 LOCATION: 1..4707
 US-09-459-065-1

Query Match 3.5%; Score 201.2; DB 4; Length 4707;
 Best Local Similarity 49.1%; Pred. No. 4.3e-39;
 Matches 824; Conservative 0; Mismatches 728; Indels 126; Gaps 6;

Qy 3935 TTCGCTTCTGAGATTAATGAACTTACACCATGCGGAGTGTGTTCCAGGAATTAAGACC 3994
 Db 3038 TTCCCTTCCTTGACTACAAACATTTTGTGCTGAGAACTTCTTCCAGATCAGGTGCT 3097
 Qy 3995 ACCCTGCTCCGGAGCCTTGAAGTCCCGGCTACCGGAGAGAGCTGTGAGAAAGGCC 4054
 Db 3098 TCACCCATCTTCACTGAATATGATTAACAGAGACGCCAAGACAGATGAAGTTC 3157
 Qy 4055 TGAAGCTCTTGCCAGCTCATCAACAAAGAGTGTCTGCTGCTTCTTATCCGACGC 4114
 Db 3158 TCACAGCTTTGAGATGCTTATCTGTATTAAGATTTCTTGTATGATTCACACACC 3217
 Qy 4115 TTGAGTCCAGCGTAGCTTCTCATGCGCGACCGTGGACAGTGGCTCACTCATGA 4174
 Db 3218 TTGAAGACAGAAAGATTTTCTGTGAAGGACAGGTGTGTTGCTTCTTCTTAACA 3277
 Qy 4175 CCGTGTCCAGAGAGAGCTGAGTACGCCCATGATGTGCTGAGCAGCTGCGCCGACC 4234
 Db 3278 TTGACCTCAACCAAGTGTCTTACCTGACCAAGATCTTGAAGTGTGACAGGAGCT 3337
 Qy 4235 TCATTGACAGAACCTGAGAGCAAGAACCAACCTTAAGCTGTCTCAGAGACTGAGT 4294
 Db 3338 TGAT-----GAAACAGTGTGTAAACATGACGCCGAAATCATGCTGAGACAGGAGT 3391
 Qy 4295 CAGTGTGAGAGAGATGCTGACCAATGCTTACCTTCTGCTCAAGTTCCTCAAG 4354
 Db 3392 CCGTGTGAGAAACTCTCAAACTGATGATGATGATGATGATGATGATGATGATGATG 3451
 Qy 4355 AGTGTGTGGGAGGCCCTCTTCTCCCTGCTGTGCTGCTGCTGCTGCTGCTGCTGCT 4414
 Db 3452 AGCTGTGAGAGAGCCTTCTTATTTGCTGTGAGCAGCTGTAACCAAGAAATTAACA 3511
 Qy 4415 GCCCATTTGAGCGCATACCGGCGAGGCCCTCTCTCTTGAAGGAGGACAGCTATCC 4474
 Db 3512 GTCCGCTGATGATTAATCACTTGAAGGCCCTGTACACTTAATGAACCTGGCTGTT 3571
 Qy 4475 GCCAGCATTTGACTACAAACCCCTGCTCTGAGCTGTGAGCCGACAGCAATGCCACA 4534
 Db 3572 GGCAGGTTCCGGAATTCAGTACTGTGCAATTAAGTGTCTTTGAAGAAATCCCGAAA 3631
 Qy 4535 GC-----CCGAGTCCAGTAAAGATCTCAATCTCAATCTCAATCTCAATCTCA 4576
 Db 3632 ACAGAGTGAGAGATGCTGTGGAATATTTCAATGATTTCTGACTGTGACACATTTG 3691
 Qy 4577 CTCAGGTCAAGAGAAAGATTCGTGATGCCATCTTCAAGAAATGCTGCTTCCACCGGC 4636
 Db 3692 GCCAGCCAAAGAAAGATTTTCCAGATTTCTTAAGAAATTAAGGCTCTCTTAATGAC 3751
 Qy 4637 CCAAAGCTGAGATATGATGATGATGATGATGATGATGATGATGATGATGATGAT 4696
 Db 3752 TTCACTTAATGAATTTGCTGTGAGCTTCAATGAGGACACAGACAGAAAGACTTGG 3811
 Qy 4697 AGAGTGAAGACATCAACCAAGATTTGATGATGATGATGATGATGATGATGATGAT 4756
 Db 3812 ACATGACAGTCTTCTCGTATTTCTTGAATGATGATGATGATGATGATGATGATGAT 3871
 Qy 4757 ACTACAGGTGCAAGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 4816
 Db 3872 ACTATGAGATATCAATGATGATGATGATGATGATGATGATGATGATGATGATGAT 3931
 Qy 4817 ACGCAGTGAACATCCACCGCTCTCAGGACCTCAGCAAGTAAATATGAATGAATGAT 4876

```

Db      3932 CAGATGTGAGTACTTC-----GGATGACCATGTGCATTGTATTTTA----- 3972
Qy      4877 GGTACAGGGGACGCCCGACAGCTCCGGCTCAGGACACCTATGATCACTCTGACCTGG 4936
Db      3973 ----- 3972
Qy      4937 AGAGTGAAGTCAAGATGTGSCACCTAGTGAAGAACACGAGCACGAGACCAAGAGAGG 4996
Db      3973 -----CCAGATTGGAGAAATTCACAGATGTGCAGAGAAAGAGAC-----ATC 4015
Qy      4997 GGGAGCCGGGGAGGAGAGTGTGTCTGAAATCTACCTGACCCGACTCCTGGCCACTAAG 5056
Db      4016 GAGGAGACACAACTTCAAAGTAAAGAAATGTATCTTGACAAAGCTGTCTGACCAAGG 4075
Qy      5057 GCACACTGCAGAAAGTTGTGTGATGACCTCTTTGAGACCATTTTGACACGGCACACCTGG 5116
Db      4076 TGGCAATTCATTCTGTGTCTGTAATAAATTGTAAGACATTTGAGATTGCCAACAGCA 4135
Qy      5117 GCTCTGCCCTGCCCCCTGACCATCAAGTACATGTTTGAATCTTCTGATGAGAGAGCTGATA 5176
Db      4136 GAGCT-----CCATTGCTATTAATATCTTTTGTGACTTTTGGACGCCCGAGCTGAAA 4189
Qy      5177 AACATGGAATTCATGACCCGACGTCGCGCATACCTGGAAGAGCAATTTGCTGCCCTGA 5236
Db      4190 ACAAAATAACACAGATCCTGACGTCGTACATATTTGAAACAAACAGCTTCCTCTTC 4249
Qy      5237 GGTTTGGGTCAACATGATCAAGAACCCGCACTTTGTGTGTGACATCCATAGAACAGCA 5296
Db      4250 GCTTCTGGGTAAACATCTCGAAGAACCTCGATTTGTGTGACATTAAGAAACACACAC 4309
Qy      5297 TCACAGACGCTGCTCTCTGTGTGTGCTCAGACCTTCATGAGACTTTGCTCCACGTCAG 5356
Db      4310 ATATGAGCGGCTGTTTGTTCAGATGATGGCCAGGCACTTATGAGATGATTTTCTCTCAG 4369
Qy      5357 AGCACCGGCTGGGCAAGACCTGCCCCCAACAGCTGTGTATGCCAAGACATCTCCCA 5416
Db      4370 AGCACCACTAGGGAAGGAAGACCACTAATTAAGCTCTCTATGCCAAGATATCCCA 4429
Qy      5417 GCTACAGAAATGGGTGAGAGATTTACTCAGACATAGGGAAGATGCCACCATCAAGC 5476
Db      4430 CTTACAGAAAGAAATTAATTTATTCAGAAAGCAATCAGGATTTGCTCCTCATGTGAT 4489
Qy      5477 ACCAAGACATGAACGATACCTGGCTGAGCAGTCCCGAGTGACATGATAGTTCAACA 5536
Db      4490 CCTCAGAAATGGAAGATTTTAACTCAGGAATCTAGAAACATGAAATGATTTATAG 4549
Qy      5537 CCATAGTGCATCTCAGAGATCTTCTCTATGTGGGAAATAGACGAGAGATCTT 5594
Db      4550 AAGAACTGCTTGACAGAAATTTCAATAATCATGTAATAATTTTATGATGATTC 4607

```

```

RESULT 9
US-09-328-475C-37/c
; Sequence 37, Application US/09328475C
; Patent No. 6476207
; GENERAL INFORMATION:
; APPLICANT: Zhang, Jimmy
; APPLICANT: Astel, Jon H.
; APPLICANT: Carroll III, Eddie
; APPLICANT: Endege, Wilson O.
; APPLICANT: Ford, Donna M.
; APPLICANT: Monahan, John E.
; APPLICANT: Schlegel, Robert
; APPLICANT: Steinmann, Kathleen E.
; TITLE OF INVENTION: GENES AND GENE EXPRESSION PRODUCTS THAT
; FILE REFERENCE: 1532.002/200130.463
; CURRENT APPLICATION NUMBER: US/09/328,475C
; NUMBER OF SEQ ID NOS: 341
; SOFTWARE: PatsSeq for Windows Version 3.0
; SEQ ID NO 37

```

```

; LENGTH: 1024
; TYPE: DNA
; ORGANISM: Homo Sapien
; FEATURE:
; NAME/KEY: misc.feature
; LOCATION: (1)..(1024)
; OTHER INFORMATION: n = A,T,C or G
US-09-328-475C-37
Query Match 1.5%; Score 86.4; DB 4; Length 1024;
Best Local Similarity 67.9%; Pred. No. 2,4e-11;
Matches 148; Conservative 0; Mismatches 68; Indels 2; Gaps 2;
Qy      5472 CAGCGACCAAGACATGAAGCATAT-CCTGGCTGAGCACTCCCGATGCAATGAATGAG- 5529
Db      780 CAGTACACAGGCCATGAATGCTTACCCTGCGAGCAGTCCCGCTGCACNCCGTGAGT 721
Qy      5530 TTCAACACCATAGATGACACTTCAAGATCTTCTCTATGTGGGAATTAACGACGAG 5589
Db      720 TTCAACATGTTGAGTGGCCCGCCATGATCTACTCTATGTGCAAGTATGTAGAG 661
Qy      5590 ATCCTTGAACCTCTGACACCATGATGACAGAGTGTGGAAAGCAAACTGGCTTACAACTA 5649
Db      660 CTGATCGGGGCCCTTAGAGAGATGAGAGAGGACGCGCGCANCGCTTGCTTATAAGTG 601
Qy      5650 GAACAGTCATTAACCTCATGAGCTTTAGACAGCTGAAA 5687
Db      600 GAGCAGCTCATTAATGCCATGTCATTGAGAGCTGAGA 563

```

```

RESULT 10
US-09-313-294A-6281
; Sequence 6281, Application US/09313294A
; Patent No. 6476212
; GENERAL INFORMATION:
; APPLICANT: Lalzudi, Raghunath V.
; APPLICANT: Ito, Laura Y.
; APPLICANT: Sherman, Bradley K.
; TITLE OF INVENTION: POLYNUCLEOTIDES AND POLYPEPTIDES DERIVED FROM CORN EAR
; FILE REFERENCE: PL-0017 US
; CURRENT APPLICATION NUMBER: US/09/313,294A
; NUMBER OF SEQ ID NOS: 7600
; SOFTWARE: PERL Program
; SEQ ID NO 6281
; LENGTH: 288
; TYPE: DNA
; ORGANISM: Zea mays
; FEATURE:
; NAME/KEY: misc.feature
; OTHER INFORMATION: Incyte ID No. 6476212 700351536H1
; LOCATION: 9, 19, 26, 50, 82, 109, 135, 175-176, 186, 197, 208, 217, 244
; OTHER INFORMATION: a, t, c, g, or other
US-09-313-294A-6281
Query Match 1.0%; Score 57.8; DB 4; Length 288;
Best Local Similarity 73.2%; Pred. No. 0.00013;
Matches 71; Conservative 0; Mismatches 26; Indels 0; Gaps 0;
Qy      4685 GGATGATCTTGCAGATGAAGACATACACCAAGATTGAGATGATTTGGAAGCACTGA 4744
Db      12 GGGTCGTCGCAANTGAGAGACATCACCAAGATTNAGGTGACTGGAAAGCGCTCA 71
Qy      4745 ACAACCTGGCCCACTACCAAGTCCAGATGTTCCGT 4781
Db      72 ACAACCTGATNCATTTACAGTGAAGGGGGGTCTTT 108

```

```

RESULT 11
US-08-232-463-14/c
; Sequence 14, Application US/08232463
; Patent No. 5670367

```

GENERAL INFORMATION:
APPLICANT: DOENER, F.
APPLICANT: SCHEITFLINGER, F.
APPLICANT: FALKNER, F. G.
TITLE OF INVENTION: RECOMBINANT FOWLPOX VIRUS
NUMBER OF SEQUENCES: 52
CORRESPONDENCE ADDRESS:
ADDRESSEE: Foley & Lardner
STREET: 1800 Diagonal Road, Suite 500
City: Alexandria
STATE: VA
COUNTRY: USA
ZIP: 22313-0299
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/232,463
FILING DATE:
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US/07/935,313
FILING DATE:
APPLICATION NUMBER: EP 91 114 300.6
FILING DATE: 26-AUG-1991
ATTORNEY/AGENT INFORMATION:
NAME: BENT, Stephen A.
REGISTRATION NUMBER: 29,768
REFERENCE/DOCKET NUMBER: 30472/114 IMMU
TELECOMMUNICATION INFORMATION:
TELEPHONE: (703)836-9300
TELEFAX: (703)683-4109
TELEX: 899149
INFORMATION FOR SEQ ID NO: 14:
SEQUENCE CHARACTERISTICS:
LENGTH: 7218 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
IMMEDIATE SOURCE:
CLONE: PTZpT-Fls
US-08-232-463-14

Query Match 0.9%; Score 49.6; DB 1; Length 7218;
Best Local Similarity 4.0%; Pred. No. 0.058; Mismatches 163; Indels 0; Gaps 0;
Matches 16; Conservative 219;

4649 ATATGATCTGATGCGCAAGAGTGGGCAAGATGATCTTGACAGATGAAGACA 4708
1443 ATTGGACRR 1384
4709 TCACCAACAAGATTGAGATTTGGAAGCATGAACACACTGCCCACTACCAAGTGC 4768
1383 RRR 1324
4769 CAGATGTCCTGCTGATGATGTCACAGAGTGAAGCCTTAACGACAGTGAACA 4828
1323 RRR 1264
4829 ACTCCACGCTCCAGACCTCAGCAAGTAATATGAAAATGATCCGTAACGGGCA 4888
1263 RRR 1204
4889 GCCCGCAGCGCTCCGCTCAGCAGACCTATGATCACTCTCTGAGAGTGAAGTCA 4948
1203 RRR 1144
4949 AGATGTGACCTAGTGAAGACAGACGACGACCAAGAGAGGAGGAGCCGGGGA 5008
1143 RRR 1084

5009 GCAAGATGCTGTGAATCTTACCTGACCCGACTCTG 5046
1083 RRR 1046

RESULT 12
US-08-607-509-3
Sequence 3, Application US/08607509
Patent No. 5876735
GENERAL INFORMATION:
APPLICANT: Reed, Steven G.
TITLE OF INVENTION: METHODS FOR ENHANCEMENT OF PROTECTIVE IMMUNE RESPONSES
NUMBER OF SEQUENCES: 15
CORRESPONDENCE ADDRESS:
ADDRESSEE: Seed and Berry LLP
STREET: 6300 Columbia Center, 701 Fifth Avenue
City: Seattle
STATE: Washington
COUNTRY: USA
ZIP: 98104-7092
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/607,509
FILING DATE: 16-FEB-1996
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: MAKI, David J.
REGISTRATION NUMBER: 31,392
REFERENCE/DOCKET NUMBER: 210121.404C3
TELECOMMUNICATION INFORMATION:
TELEPHONE: (206) 622-4800
TELEFAX: (206) 682-6031
INFORMATION FOR SEQ ID NO: 3:
SEQUENCE CHARACTERISTICS:
LENGTH: 1867 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
FEATURES:
NAME/KEY: CDS
LOCATION: 117..1325
US-08-607-509-3

Query Match 0.9%; Score 49.4; DB 2; Length 1867;
Best Local Similarity 45.3%; Pred. No. 0.034; Mismatches 216; Indels 0; Gaps 0;
Matches 179; Conservative 0;

4049 AAGGCTGAAGCTTTGCGCCAGCTCATCAACAAGAGTGTCTCTGCTTATCC 4108
403 AGGGCTGTGCTCTCCCCCACTGCGAGCTGGCCCTGCAAGCGAGGTGATCAGCC 462
4109 GCAGCTTGAATCCAGAGTGTCTTCCCATGCGGACCGGACCGGACGCTCACTCA 4168
463 GCATGGTGAATCTCTGCAACAGCTTCAAGTTCTGCGAAGACTTTGTGCGCGGACGC 522
4169 TCATGACCGTCTGCAAGAGAGTGAAGTGAAGCACTGATGTGCTGAAGCAGTGTG 4228
523 GCGTCAGAGATGACCTGCGAGAGTGAAGCGCGCGCTCATGCTGCGTGCGACGCGCG 582
4229 CCGACTATTTGACAAAGAACTGGAAGCAAGAACCCCTAAGCTGTGCTGCGAGAGA 4288
583 GCGGCTGTCCGAGCTGATCAAGGTGGCGCTGCGACAGATCGCTGCGCGTGTG 642
4289 CTGATCAGTGTGCTGAGAGATGCTGACCAATGCTTACTTCTCTCTACAAAGTTCC 4348
643 TGCTGAGAGAGCTGATGATGCTGTCTGAGGCTTTCGCGACAGATTTAACAGATCT 702
4349 TCAAGAGATGTGCTGGAGAGCCCTTTCTCTCTGCTTCTGTGCGCATCAAGAGAGATG 4408

Mon Feb 23 09:34:02 2004

us-09-964-956-12.rn1

Page 12

Db 703 TCCGCTTCTGCGGAGGACATCCAGTGGCGCTTCTCGGCCACGATGCCGAGAGAG 762
Qy 4409 AGAAGGCCCCCATTGACCCCATCAGCGGAGGCC 4443
Db 763 TACTGAGCTGACGAAGAAGTTCAATGCGGACCCC 797

RESULT 13

US-08-634-642-3

Sequence 3, Application US/08634642

Patent No. 5879687

GENERAL INFORMATION:

APPLICANT: Reed, Steven G.

TITLE OF INVENTION: METHODS FOR ENHANCEMENT OF

TITLE OF INVENTION: PROTECTIVE IMMUNE RESPONSES

NUMBER OF SEQUENCES: 15

CORRESPONDENCE ADDRESS:

ADDRESSEE: SEED and BERRY LLP

STREET: 6300 Columbia Center, 701 Fifth Avenue

CITY: Seattle

STATE: Washington

COUNTRY: USA

ZIP: 98104-7092

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patent in Release #1.0, Version #1.30

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/634,642

FILING DATE: 18-Apr-1996

CLASSIFICATION: 424

ATTORNEY/AGENT INFORMATION:

NAME: Mark, David J.

REGISTRATION NUMBER: 31,392

REFERENCE/DOCKET NUMBER: 210121.404C4

TELECOMMUNICATION INFORMATION:

TELEPHONE: (206) 682-4900

TELEFAX: (206) 682-6031

INFORMATION FOR SEQ ID NO: 3:

SEQUENCE CHARACTERISTICS:

LENGTH: 1867 base pairs

TYPE: nucleic acid

STRANDEDNESS: single

TOPOLOGY: linear

FEATURE:

NAME/KEY: CDS

LOCATION: 117..1325

US-08-634-642-3

Query Match 0.9%; Score 49.4; DB 2; Length 1867;
Best Local Similarity 45.3%; Pred. No. 0.034;
Matches 179; Conservative 0; Mismatches 216; Indels 0; Gaps 0;

Qy 4049 AAGGCTTGAAGCTTTGGCCAGCTCATCAACAAGGTGTTCTGCTGTCTTCATCC 4108
Db 403 AGGCGCTGCTGCTCTCCCACTGCGAGCTGCGCCCTGCAAGCGGAGGTATGAGCC 462
Qy 4109 GCAAGCTTGAAGTCCCAAGTATGTTTCCATGCGCGACCGTGGCAAGTGGCTCACTCA 4168
Db 463 GCATCGGTGAGTCTCTTCGAAACAGCTCAAGTTTCTGCGAAGCTTTTGTGGGCGACGC 522
Qy 4169 TCATGACCGTCTGCTGCAAGAGTGTGATGCGCACTGATGTGCTGAAGCAAGTCTGG 4228
Db 523 GCGTGCAGAGATGACTGTGGCAAGTGTGCAAGCGCGCGTCTGCTGCGGCGACCGCG 582
Qy 4229 CCGACCTTATGACAAAGACCTGAGAGCAAGCAACCTTAAGTGTGCTGCTGAGAGGA 4288
Db 583 GCGCGTGTCCGAGCTGATCAAGGCTGCGCGCTGCGCAAGAGTCCGCGCTGCTGG 642
Qy 4289 CTGAGTCAAGTGTGAGAGATGTGACCAATTGTTTACTTCTCTCTTACAAATTC 4348
Db 643 TGCTGACAGAGGCTGATGATGTGCTGCTGAGGCTTGGGAGCAAGATTTACGAGATCT 702

Qy 4349 TCAAGAGTGTGTGAGGAGACCCCTTCTCTCCCTGTCTGTGCTCAAGACAGATAG 4408
Db 703 TCCGCTTCTGCGGAGGACATCCAGTGGCGCTTCTTCGCGCACGATGCCGAGAGAG 762
Qy 4409 AGAAGGCCCCCATTGACCCCATCAGCGGAGGCC 4443
Db 763 TACTGAGCTGACGAAGAAGTTCAATGCGGACCCC 797

RESULT 14

US-08-989-370-3

Sequence 3, Application US/08989370

Patent No. 6013268

GENERAL INFORMATION:

APPLICANT: Reed, Steven G.

TITLE OF INVENTION: METHODS FOR ENHANCEMENT OF PROTECTIVE IMMUNE RESPONSES

NUMBER OF SEQUENCES: 14

CORRESPONDENCE ADDRESS:

ADDRESSEE: SEED and BERRY

STREET: 6300 Columbia Center, 701 Fifth Avenue

CITY: Seattle

STATE: Washington

COUNTRY: USA

ZIP: 98104-7092

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patent in Release #1.0, Version #1.30

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/989,370

FILING DATE: 12-DEC-1997

CLASSIFICATION:

ATTORNEY/AGENT INFORMATION:

NAME: Mark, David J.

REGISTRATION NUMBER: 31,392

REFERENCE/DOCKET NUMBER: 210121.404C5

TELECOMMUNICATION INFORMATION:

TELEPHONE: (206) 622-4900

TELEFAX: (206) 682-6031

INFORMATION FOR SEQ ID NO: 3:

SEQUENCE CHARACTERISTICS:

LENGTH: 1867 base pairs

TYPE: nucleic acid

STRANDEDNESS: single

TOPOLOGY: linear

FEATURE:

NAME/KEY: CDS

LOCATION: 117..1325

US-08-989-370-3

Query Match 0.9%; Score 49.4; DB 3; Length 1867;
Best Local Similarity 45.3%; Pred. No. 0.034;
Matches 179; Conservative 0; Mismatches 216; Indels 0; Gaps 0;

Qy 4049 AAGGCTTGAAGCTTTGGCCAGCTCATCAACAAGGTGTTCTGCTGTCTTCATCC 4108
Db 403 AGGCGCTGCTGCTCTCCCACTGCGAGCTGCGCCCTGCAAGCGGAGGTATGAGCC 462
Qy 4109 GCAAGCTTGAAGTCCCAAGTATGTTTCCATGCGCGACCGTGGCAAGTGGCTCACTCA 4168
Db 463 GCATCGGTGAGTCTCTTCGAAACAGCTCAAGTTTCTGCGAAGCTTTTGTGGGCGACGC 522
Qy 4169 TCATGACCGTCTGCTGCAAGAGTGTGATGCGCACTGATGTGCTGAAGCAAGTCTGG 4228
Db 523 GCGTGCAGAGATGACTGTGGCAAGTGTGCAAGCGCGCGTCTGCTGCGGCGACCGCG 582
Qy 4229 CCGACCTTATGACAAAGACCTGAGAGCAAGCAACCTTAAGTGTGCTGCTGAGAGGA 4288
Db 583 GCGCGTGTCCGAGCTGATCAAGGCTGCGCGCTGCGCAAGAGTCCGCGCTGCTGG 642
Qy 4289 CTGAGTCAAGTGTGAGAGATGTGACCAATTGTTTACTTCTCTCTTACAAATTC 4348

Mon Feb 23 09:34:02 2004

us-09-964-956-12.rn1

Page 13

Db 643 TGCTGACAGGCTGATGATGCTCTCTCAGGGCTTCGCGACCAAGATTTCAGAGATCT 702
QY 4349 TCAGAGAGTGTCTGGGGGAGCCCTCTTCTCCCTGTTCTGTGCATCAAGCAGCATG 4408
Db 703 TCCGCTTCTGCGAAGACATTCAGTTCGCTCTTCTCCCGCAGATTCGGAGAGG 762
QY 4409 AGAAGGCCCCATTGACGCCATCACGGCGAGGCC 4443
Db 763 TACTGAGCTGACGAAGAAGTTTCATGCGGACCCC 797

RESULT 15
US-09-398-169-3
; Sequence 3, Application US/09398169
; Patent No. 6660840

GENERAL INFORMATION:

APPLICANT: Reed, Steven G.
TITLE OF INVENTION: METHODS FOR ENHANCEMENT OF PROTECTIVE IMMUNE RESPONSES
NUMBER OF SEQUENCES: 14
CORRESPONDENCE ADDRESS:
ADDRESSEE: SEED and BERRY
STREET: 6300 Columbia Center, 701 Fifth Avenue
City: Seattle
STATE: Washington
COUNTRY: USA
ZIP: 98104-7092

COMPUTER READABLE FORM:

MEDIUM TYPE: floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent In Release #1.0, Version #1.30

CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/398,169

FILING DATE: 17-Sep-1999
CLASSIFICATION: <Unknown>

PRIOR APPLICATION DATA:
APPLICATION NUMBER: US/08/989,370

FILING DATE: 12-DEC-1997
ATTORNEY/AGENT INFORMATION:

NAME: Maki, David J.
REGISTRATION NUMBER: 31,392

REFERENCE/DOCKET NUMBER: 210121.404CS
TELECOMMUNICATION INFORMATION:

TELEPHONE: (206) 622-4900
TELEFAX: (206) 682-6031

INFORMATION FOR SEQ ID NO: 3:

SEQUENCE CHARACTERISTICS:
LENGTH: 1867 base pairs

TYPE: nucleic acid
STRANDEDNESS: single

TOPOLOGY: linear

FEATURE:
NAME/KEY: CDS
LOCATION: 117..1325

SEQUENCE DESCRIPTION: SEQ ID NO: 3:
US-09-398-169-3

Query Match 0.9%; Score 49.4; DB 4; Length 1867;
Best Local Similarity 45.3%; Pred. No. 0.034;
Matches 179; Conservative 0; Mismatches 216; Indels 0; Gaps 0;

QY 4049 AAGGCTGAAGCTTTGGCCAGCTCATCAACAAGAGTTCCTCTCTCTATCC 4108
Db 403 AGGGCTCTGCTCTCCCGCACTGCGAGCTGCGCAGAGGCGAGGATATCAGCC 462

QY 4109 GCAGCTGAGTCCGAGGCTTCTTCATGCGGACCGTGGCAAGTGGCTCACTCA 4168
Db 463 GCATGGGATTTCTGTGCAACAGCTCAAGTTCTGAGACCTTTGTGGGGGACGCG 522

QY 4169 TCATGACCGTGTGCAAGCAAGCTGAGTACGCCACTGATGTCTGAAGAGCTGTGG 4228
Db 523 GCGTGACAGATGACCTGCGCAAGCTGCGAGCGCGGCTTCATGCTTGCCTGAGGCAAGCGCGG 582

QY 4229 CCGACTCATTTGACMAAGACCTGAGAGCAAGAACCACTTAAGCTGTCTCAGAGGA 4288
Db 583 GCCCGGTGTCCGACGTGATCAAGCGTGGCCGCTGCGCAAGAGTGGCTGGCTGG 642
QY 4289 CTGAGTCAAGGCTGAGAGATGCTGACCAATTGTTTACTTCTCTCTACAAATTC 4348
Db 643 TGCTGACGAGGCTGATGATGATCTGTCTGAGGCTTCGCGACCAAGATTTCAGAGATCT 702
QY 4349 TCAGAGAGTGTCTGGGAGCCCTCTTCTCTCTGTCTGTGCATCAAGCAGCATG 4408
Db 703 TCCGCTTCTGCGAAGACATCCAGGTCCGCTCTTCTCCCGCAGATGCCGAGAGG 762
QY 4409 AGAAGGCCCCATTGACGCCATCACGGCGAGGCC 4443
Db 763 TACTGAGCTGACGAAGAAGTTTCATGCGGACCCC 797

Search completed: February 19, 2004, 23:18:28
Job time : 272 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: February 19, 2004, 20:48:15 ; Search time 1236 Seconds
(without alignments)
16122.863 Million cell updates/sec

Title: US-09-964-956-12

Perfect score: 5691

Sequence: 1 atgaagcctcgcctcgga.....gctagacagctgaataaa 5691

Scoring table: IDENTITY_NUC

Gapop 10.0 , Gapext 1.0

Searched: 2308684 seqs, 1750822206 residues

Total number of hits satisfying chosen parameters: 4617368

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Published Applications_NA:*

1: /cgn2_6/ptodata/1/pubpna/US07_PUBCOMB.seq:*
2: /cgn2_6/ptodata/1/pubpna/PCT_NEM_PUB.seq:*
3: /cgn2_6/ptodata/1/pubpna/US06_NEM_PUB.seq:*
4: /cgn2_6/ptodata/1/pubpna/US06_PUBCOMB.seq:*
5: /cgn2_6/ptodata/1/pubpna/US07_NEM_PUB.seq:*
6: /cgn2_6/ptodata/1/pubpna/PCTUS_PUBCOMB.seq:*
7: /cgn2_6/ptodata/1/pubpna/US08_NEM_PUB.seq:*
8: /cgn2_6/ptodata/1/pubpna/US08_PUBCOMB.seq:*
9: /cgn2_6/ptodata/1/pubpna/US09_PUBCOMB.seq:*
10: /cgn2_6/ptodata/1/pubpna/US09_PUBCOMB.seq:*
11: /cgn2_6/ptodata/1/pubpna/US09_NEM_PUB.seq:*
12: /cgn2_6/ptodata/1/pubpna/US10_PUBCOMB.seq:*
13: /cgn2_6/ptodata/1/pubpna/US10_PUBCOMB.seq:*
14: /cgn2_6/ptodata/1/pubpna/US10_PUBCOMB.seq:*
15: /cgn2_6/ptodata/1/pubpna/US10_PUBCOMB.seq:*
16: /cgn2_6/ptodata/1/pubpna/US10_PUBCOMB.seq:*
17: /cgn2_6/ptodata/1/pubpna/US60_NEM_PUB.seq:*
18: /cgn2_6/ptodata/1/pubpna/US60_PUBCOMB.seq:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match Length	ID	Description
1	2778.4	48.8	6730	US-10-175-523-95
2	2533.8	44.5	5895	US-10-087-684-31
3	2533.8	44.5	5895	US-10-218-779-31
4	1157.8	20.3	3666	US-10-108-260A-802
5	793.8	13.9	2597	US-10-245-103-91
6	793.8	13.9	2597	US-10-245-107-91
7	793.8	13.9	2597	US-10-245-143-91
8	793.8	13.9	2597	US-10-245-171-91
9	793.8	13.9	2597	US-10-245-851-91
10	793.8	13.9	2597	US-10-245-883-91
11	793.8	13.9	2597	US-10-237-535-91
12	793.8	13.9	2597	US-10-238-183-91
13	793.8	13.9	2597	US-10-238-283-91
14	793.8	13.9	2597	US-10-238-370-91
15	793.8	13.9	2597	US-10-245-055-91

16	793.8	13.9	2597	US-10-245-147-91	Sequence 91, Appl
17	793.8	13.9	2597	US-10-245-730-91	Sequence 91, Appl
18	793.8	13.9	2597	US-10-245-739-91	Sequence 91, Appl
19	793.8	13.9	2597	US-10-246-210-91	Sequence 91, Appl
20	793.8	13.9	2597	US-10-239-196-91	Sequence 91, Appl
21	793.8	13.9	2597	US-10-243-024-91	Sequence 91, Appl
22	793.8	13.9	2597	US-10-243-409-91	Sequence 91, Appl
23	793.8	13.9	2597	US-10-245-821-91	Sequence 91, Appl
24	793.8	13.9	2597	US-10-245-880-91	Sequence 91, Appl
25	793.8	13.9	2597	US-10-245-033-91	Sequence 91, Appl
26	793.8	13.9	2597	US-10-243-095-91	Sequence 91, Appl
27	793.8	13.9	2597	US-10-245-185-91	Sequence 91, Appl
28	793.8	13.9	2597	US-10-245-427-91	Sequence 91, Appl
29	793.8	13.9	2597	US-10-245-473-91	Sequence 91, Appl
30	793.8	13.9	2597	US-10-245-770-91	Sequence 91, Appl
31	793.8	13.9	2597	US-10-245-877-91	Sequence 91, Appl
32	793.8	13.9	2597	US-10-246-576-91	Sequence 91, Appl
33	793.8	13.9	2597	US-10-243-520-91	Sequence 91, Appl
34	793.8	13.9	2597	US-10-242-743-91	Sequence 91, Appl
35	793.8	13.9	2597	US-10-242-845-91	Sequence 91, Appl
36	793.8	13.9	2597	US-10-237-636-91	Sequence 91, Appl
37	793.8	13.9	2597	US-10-238-325-91	Sequence 91, Appl
38	793.8	13.9	2597	US-10-238-346-91	Sequence 91, Appl
39	793.8	13.9	2597	US-10-238-411-91	Sequence 91, Appl
40	793.8	13.9	2597	US-10-243-424-91	Sequence 91, Appl
41	793.8	13.9	2597	US-10-243-445-91	Sequence 91, Appl
42	793.8	13.9	2597	US-10-243-446-91	Sequence 91, Appl
43	793.8	13.9	2597	US-10-245-874-91	Sequence 91, Appl
44	793.8	13.9	2597	US-10-242-653-91	Sequence 91, Appl
45	793.8	13.9	2597	US-10-243-167-91	Sequence 91, Appl

ALIGNMENTS

RESULT 1
US-10-175-523-95
Sequence 95, Application US/10175523
Publication No. US20030096264A1
GENERAL INFORMATION:
APPLICANT: Brockman, Jeffrey
APPLICANT: Evans, David
APPLICANT: Hook, Derek
APPLICANT: Klimczak, Leszek
APPLICANT: Laeng, Pascal
APPLICANT: Palfreyman, Michael
TITLE OF INVENTION: MULTI-PARAMETER HIGH THROUGHPUT SCREENING ASSAYS (MPHTS)
FILE REFERENCE: 3235/11795-US3
CURRENT APPLICATION NUMBER: US/10/175,523
CURRENT FILING DATE: 2002-06-18
PRIOR APPLICATION NUMBER: US 60/299,151
PRIOR FILING DATE: 2001-06-18
PRIOR APPLICATION NUMBER: US 60/317,828
PRIOR FILING DATE: 2001-09-07
PRIOR APPLICATION NUMBER: US 60/325,150
PRIOR FILING DATE: 2001-09-25
PRIOR APPLICATION NUMBER: US 60/333,047
PRIOR FILING DATE: 2001-11-14
PRIOR APPLICATION NUMBER: US 60/349,936
PRIOR FILING DATE: 2002-01-18
PRIOR APPLICATION NUMBER: US 60/361,834
PRIOR FILING DATE: 2002-03-04
NUMBER OF SEQ ID NOS: 197
SOFTWARE: PatentIn version 3.1
SEQ ID NO 95
LENGTH: 6730
TYPE: DNA
ORGANISM: Mus musculus
FEATURE:
NAME/KEY: misc feature
LOCATION: (1)-(6730)
OTHER INFORMATION: where n may be a or g or c or t/u, unknown, or other

US-10-175-523-95

Query Match 48.8%; Score 2778.4; DB 14; Length 6730;
Best Local Similarity 69.3%; Pred. No. 0;
Matches 3858; Conservative 0; Mismatches 1691; Indels 15; Gaps 5;

QY 137 AGCCCGCCGAGGTTTCATACCTGTTGATGAGAGACAGACACATTACTTGG 196
DB 683 ATGTGACTGACCTTTCAACATTGACTGACACGAGACAGGGGCTGTGTATGG 742
QY 197 GGGCCCTGATCGGATTTTACAGCTCTCCAGCCGCTGAAGCTTTGTGACGCAAGA 256
DB 743 GGGGTATCATGCTGTCTACAGTTGACTGGCAACCTCCATCCAGGTGCTCAAGG 802
QY 257 CAGGGCCGAGAGAGACACCCCAAGTTTACCACCCCGCATGCTCAGACCTGCAATG 316
DB 803 CAGGGCCGAGAGAGACACAGGCTTGTACCCACCTCATTTGTACAGCTCTGAGTG 862
QY 317 AGCCCTGACACCAACCAATGTCAACAGATGCTCTCTATGATTAAGAGACCA 376
DB 863 AAGTCTTACACTCAACCAACATGTCAACAACTACGATCATTTGACTCTGAGATC 922
QY 377 GAGTATGCTCTGGGAGCCTGTACCAAGGACATCTGCAAGCTGCTGAGGAGACC 436
DB 923 GCTGCTGCTCTGGGAGGCTCTACAGGCTGTTTCCAGCTCTGCGACTGATGAGC 982
QY 437 TCTTCAAGCTGGGAGGCTTATCATTAAGAGACATCTCTGCAAGTGTCAAGCA 496
DB 983 TCTTCAAGCTGGGAGGCTTATCATTAAGAGACATCTCTGCAAGTGTCAAGCA 1042
QY 497 GGGGCTAGCTTTGGAGTATGCTCTCTCAAGCAACCTGAGTGAACAAGTGTCTTG 556
DB 1043 CAGGACCATATATGATGTGATTTGTGGCTCTGAGGGGAGAGATGGAAGCTTTTATCG 1102
QY 557 CCACGGCAGTGAATGGAGCCCGAGTATTTTCCACATCTCCAGCCGAGAACTGACA 616
DB 1103 GCATGCTGTGATGGAGGAGGAGATTACTTCCCTACTGCTCCAGCCGAGAGTGTGCC 1162
QY 617 AGAAGCTGAGGCGAGTGGCATGTTCCGATGCTCTTCAAGATGATGTTGGGCTCGA 676
DB 1163 GTGACCTGATCTTCAAGATGCTGATGATGATGATGATGATGATGATGATGATG 1222
QY 677 TGATTAAGATCCCTCGGACACTTCAACATCATCCCTGACTTGAATCTACTATGT 736
DB 1223 TCATCAAGATCCCTCGGACACTTCAACATCATCCCTGACTTGAATCTACTATGT 1282
QY 737 ATGTTTGAAGTGGCACTTTGTCTATCTTTTGAACCTCCAGACCTGAGATGCTC 796
DB 1283 ATGTTTGAAGTGGGAGTGTGTCTATCTTCTCACTGCTCAAGATGCTGCTGCA 1342
QY 797 CACAGGCTGCA---CCACCAAGAGAGATGATATACATCCAGCTGCTGAGCTTTGCA 853
DB 1343 GCAATGGCATCAATTAACATGAGACCTTCTATACCTCAAGATGCTGCTGCA 1402
QY 854 AGAGAGACAGAGCTTCACTCTCTATGATGAGTGGCCATGGCTGAGAGGAGTGGG 913
DB 1403 AGAGAGACCCCAAGTTCACTCTCTATGCTCTGCTTTTGGCTGCAAGTGTCTGGG 1462
QY 914 TGGAGTACCGGCTGTGAGGCTGCTACTGCTCCAAAGCCGGGCGTGTGGAGGA 973
DB 1463 TGGAGTATCGCTTTTGAAGGAGCTTACTTGAACAGGAGGAGCTTACTGAGG 1522
QY 974 CCTTGAAGTCAATCAAGTGAATGATGATGCTCTTCAAGCTCTTCTCCAAAGGAGG 1033
DB 1523 CCTTGAAGTCAATCAAGGAGAGATGCTGCTTGTGCAATCTTTTCAAGGAGGAGG 1582
QY 1034 GGAATATGAATCCCTGATGATGCTGAGCTGTGATCTTCACTTTGAAGAGATTAATG 1093
DB 1583 AGTACACAGCCCTCTGATGATGCTGCTCTGCTGCTTCCCAATCCGAGCATCAACT 1642
QY 1094 ACCGCAATTAAGAGGAGGCTGAGTCTTGTATCCGAGGAGAGGAGCTGAGCTGGCT 1153
DB 1643 TGCATATCAAGAGGAGGCTGAGTCTGCTTCAAGAGAGGAGCACTTGAAGCTCAACT 1702

QY 1154 GGTCAAGTGAAGAGACATCCCTGAGCAGTGGCTCTTAACCATTTAGCATTACTTCT 1213
DB 1703 GGTGCTGTGGAAAGATGTGACAGTGAACAGAGGCGCTGTCCATGATGATTAATCTCT 1762
QY 1214 GTGGCTTGAATGAATCTCTCTGAGAGTGTCCGACATGAGTGGGAAATCCGCT 1273
DB 1763 GGGGCTGGAATCAACAGCTCTGAGGCTCTCACTCTGTGAGAGGACTGACCTGT 1822
QY 1274 TCACGAGAGACAGGAGCCGATGACATCTCTCATATGATATGCTTAAGAACCACTCTC 1333
DB 1823 ATACACAGAGAGGAGCCGCTGACCTCTGTGGCTCTATGTTTACATGGCTACAGTG 1882
QY 1334 TGGCTTTGTGGGACCAAAAGTGGCAAGCTGAAGAAATCCGGGTGATGAGACCGAGG 1393
DB 1883 TGGTTTTGTGGGACTAAGAGTGGCAAGCTGAAGAAATGAGCTGATGATGCTCCGCC 1942
QY 1394 GCAAGCCCTCAGATGAGACGCTGACAGTGT---GGAACCCCGCCAGTCTCCGGG 1450
DB 1943 ATGTGGGATCCAGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 2002
QY 1451 ATATGCTCTTCCAGAGACACAGACACTCTATATATGCTGAAGAGGAGCTCAACA 1510
DB 2003 ACATGCTCTTCCATCATGATGATGATGATGATGATGATGATGATGATGATGATGATG 2062
QY 1511 GAGTCTCTGTGGAGTCTGTGTGATGATGATGATGATGATGATGATGATGATGATGATG 1570
DB 2063 GGTCTCTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 2122
QY 1571 ACCCCACTGTGGCTGT 1630
DB 2123 ATCTCACTGT 2182
QY 1631 GGTCAAGAGAGCCCGGAGT 1690
DB 2183 GGGCTGTGGAGAAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 2242
QY 1691 ATCCCAACATATCTCGTCTCTCAAGTCAAGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1750
DB 2243 ACCCAACAGATCTCTGT 2302
QY 1751 TCCCGAGTGTCACTGT 1810
DB 2303 CTCCCACTCTCTGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 2362
QY 1811 TGGTGTGGCAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1870
DB 2363 AGGTATGTGGGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 2419
QY 1871 TCACAGAAATGGAGACCAATGCTGATGATGATGATGATGATGATGATGATGATGATG 1930
DB 2420 TCCCTGTGATGAAGCTGT 2479
QY 1931 TGACTTCCGACAGACAGCTTGTCTTCAATTTGAAGAGCTGCAATTTGCTGTGT 1990
DB 2480 AGATCTTTGTCAAGACCAATTTCAAGTTCTATTAATGAGAGTCCCAACATGCTGTGT 2539
QY 1991 CCTGCTGAGAGTGCATACCGCTGTCACTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 2050
DB 2540 CCTGTGTGAACAGCGCTTCCGCTGCAATGTTGTGAAGTGAAGTGAAGTGAAGTGAAGT 2599
QY 2051 ACCCAAGAGCTGTCTCTTCAAGAGGCGGAGTGAAGTGTGCGAGAGCTGCCCCAGC 2110
DB 2600 ACCCACTACCTGTCTCTTCAAGAGGAGAGATCAATGTTCAAGAGCTGTCTCCAGC 2659
QY 2111 TGTGTGAGTGAAGAAATCTGT 2170
DB 2660 TGTGTGCAAGAGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 2719
QY 2171 CCAAGAACTCCCGGAGCCGAGCTGTGGAGAGCTGTGATGATGATGATGATGATGATGATG 2230
DB 2720 CCGGAACTGTCCCGAGCCGAGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 2779

QY	2231	AGGGCAGGAGCAGCAGAGTGGCCCGCCGGCCCTTCAACAGCTCCAGGCTAACAGTGGCAGA	2230
Db	2780	AAGGGGCTGTCACACGGGGTCCCTGCCCTGCTTTCACAGCTTCCAGTGTGACGTGCAAA	2839
QY	2291	ACAACCTCTATTCCGTATGAAAGGATGAGATCAACCAACCTGCGGTGGAGTTGACAGTGG	2350
Db	2840	ACAGCTGGTACCAAGTATGATGGAGTGGACATCAGCAACCTAGCAGGTGACCTTCTGTAG	2899
QY	2351	TCGTGAAAGGGCACCTTCAACATTGACCAACCCAGCTCAGAAATTAAGTTACCTCTAACAT	2410
Db	2900	TATGGAATGGCAACTTCAATTATGACAAACCTCAGAGACTTGAAAGTATCATCTTACAGT	2959
QY	2411	GTGAGCCATGCGTGAAGAGCTGCGGGCTGTGCTCAGGCTGACCCAGACTTTCGATATGTG	2470
Db	2960	GTGACGCCCAAGGGGAAAGAGTGTGGTCTCTGCTCAGAGCTGACGACCAAGTTCCAGTGTG	3019
QY	2471	GCTGGTGGCAGAGGCCAGGGCCAGGTGACACCTTGGCCAGACTGTGCCCTCCAGAGAGCC	2530
Db	3020	GCTGGTGGCAGTGTGTAGGCGAGATGTACCCTTCCACAGCACTGTGCCCGACACTTCTAGCC	3079
QY	2531	AGTGGCTGAAGCTGTCTGTATGCGAAAAGCAAGTGCACAAACCCCGCATCACAGAGATTA	2590
Db	3080	CCTGGCTTGACTGTGCTCCAGCCACAAATGTCAAGTGTCCAACTCCCAATTCACAGATTT	3139
QY	2591	TCCGGGTGACAGGCCCCCGGGAAGGGGACCAAGTCACTATCCGAGGGAGAACCTGG	2650
Db	3140	TGACAGTGTACGAGACCAACCTGAAAGGAGGAGACTGTGTACCATTCATGCGCTGAACTCGG	3199
QY	2651	GCCGTGAATTTCCGACATCCGCTCCCAATGTCAGAGGTGCTGGGTGGAGTGGACGCTT	2710
Db	3200	GCTTGGACTTCTGTAGATGTGTTACCATGTGCGAGTGGCTGTGAGTGCCTTGCACACTTA	3259
QY	2711	TAGTGAATGATTACATCCCTGSCAGAACAGATCGTGTGTAGATGGGGAGGCCAACGCCA	2770
Db	3260	TCCACGGGGAAATACATCATGCTGTAGAGATGCGATGCTGTAGATGGGCCATGCCGTTATAG	3319
QY	2771	GCCAGCATGCAAGGCTTGTGTGAAGATCTGCGTGGCTGTGTGTCGGCTTAAATTCAATGGCC	2830
Db	3320	GTACCACATCTGGGGCTGTGGCGCTGTGCATTTGGGGAATGGCAAGCAAGTTCAATGACCA	3379
QY	2831	GATCCCTCAGACTTATTAATCTTCAATGACATGACTGTCTTCAGATCTGAAGCCACGCGGG	2890
Db	3380	AGTCCCAACAGCAGATATACCTTTGTGTGATCTCTGTGCTGTCACTCAGCCCGAATCGGG	3439
QY	2891	GGCCCATCTCGGAGGGACCCCAAGTGCACATACAGGGCACAACCTGATGCGGAGACA	2950
Db	3440	GACCAGATCCAGAGAGTACCAATGTGTGACCAATCACAGGCACTTACTTGTGTCTGGAGACA	3499
QY	2951	ACGTGTGTGTATGTTTGGAAAGCAGCCCTGTCTTCCACAGGGAGATCTGCATCTTCA	3010
Db	3500	GTTGTGGCAGTGACTCTGGGCAATCAGACTGTGTGAATTTAATGGGAGTCAATGAATGAGA	3559
QY	3011	TTGTGTGCAACA--CACAATCTTCAGATGAAGTGTGAGAGTGAAGATGAAGTGTGGTCAAG	3067
Db	3560	TTGTATGTGTTCACCCCACTCATCTCAATGACATGAGACCAAGTCCCTGTCCCGTAGTGTG	3619
QY	3068	TGGACAGGGCCAAATTCACACAGAGACCTGTCTTTCAGTAATGTGGAAGAACCCACACATCG	3127
Db	3620	TTCGACAGAGCCCGGGTGGATGACAGTCTGCAAGTTCAGATTAATATGATACCCACGGTCC	3679
QY	3128	TGCGGATTGACCCAGAATGAGAGCATTTGTCAGTGGAACACACCCATCGCGGTATGGGGGA	3187
Db	3680	AACGATATGACCGAGAGTGAAGTATCACTAATGGGGCACACACCTTCACTACAGGCT	3739
QY	3188	CCCACTGTGACCTCATACAGAAACCCCAAGTTCCTGTGCCAAGCATGGAAGGAGAGACACA	3247
Db	3740	TCAACTTGAATGTCAATTCAGAGGCCCAAGGTTCGAGTAATTTATGTGCAAAAGATGTG	3799
QY	3248	TCAATATCTGTGAAGTCTTGCAGACGTACTGTAGATACCTGTTCAGGCGCCGCGCTCGCTC	3307
Db	3800	TCAATGTATGACCAAGTGTGAACACACACACCCCTCACTGTGTGGCACCTCTCTGCAGACA	3855
QY	3308	TGGGTCTTGCACCAACATGACACTGACCGAGAGGCCGAGAGATTGGCTTCACTCTGG	3366

Db	3660	GTGACTACCGCCACAGCTCTGGACACTGTGGAAAGCGCCAGATTGAGTTTGGATTCTCTTTA	391.9
Oy	3368	ACAACTCCACAGTCCCTGTCTCATCTTCAACAAGACCACTTCACTACTATTCCAACCCGG	342.7
Db	3920	ACAAATGTTCAATCTTTACTCATCTATACACACAACAAGTTTCATCTACTACCCCAACCAA	397.9
Oy	3428	TGTTTGAAGCCTTTGGTCCCTCAGAGATCTCGAGAGCTCAAGCCCTGGACCCCATCATCC	348.7
Db	3580	CGTTTGAATGCTCAAGCCCACTGGATCTTTGGATCAGAAAGCCAGGCTCACCACTATCTCC	403.9
Oy	3488	TAAAGGCGAAGAACTGATCCCGCTGTGGCTGGGGCAACGTGAAGCTGATCTACACTG	354.7
Db	4040	TGAAGGCGCAAAAATCTCTGTCTCTCTCCCTCTGGAG--GGCCAAACTCACTACACAG	409.6
Oy	3548	TGCTGTGTGGGGGAAGAACCGTGACCGGTGACCGGTGAGATGTCACATGTCACCTGTCCAGT	360.7
Db	4097	TATGATTTGGAGAACACTTTTACAGTCACTGTGTCTGGACACACAGCTGTTGTGTAC	415.6
Oy	3608	CCCCCAACTCATGGCAGGACACAAGATGATGCCCCGTGTGGCTGGGATGAGTACTCC	366.7
Db	4157	CTCCCAACTCACAGGCGACAGACAAGGTCAATGTTCACTGAGGGGGAGTGGTCTTCAC	421.6
Oy	3668	CGGGAGATGATGATATTGGCCCCGACAGCCCGCTACACCTCCCGCATAGTCAGACATCG	372.7
Db	4217	CTGGCTCCGTAGAGGTATCTCCGACAGCCTGTGGACCTCTGCACAGCCATCATCTGCAATCG	427.6
Oy	3728	CAGTGCCTGGCGGCTCTCTCATCTATTTTCACTGTGCGCGTGTCAATTTGCCATTAAGCA	378.7
Db	4277	CAGCTGGATGGAAGCTCTCTTCTTATCATCTGATCACTATGTCTCATGCTTACAAAGCA	433.6
Oy	3788	AGTCCCGGAAAGTGAATCTGACCTGTAAGCGGCTGACAGATGACATGGAACAACCTGGAGT	384.7
Db	4337	AGTCTAGGAAAATGACCTCACTCAAGCGGCTCCAAATGCAATGCAACAACCTGGAGT	439.6
Oy	3848	CCCGTGTGGCCCTGAGTGTGCAAGAAACCTTTGGCCAGCTGACAGACGAGATCATGAGC	390.7
Db	4397	CCAGGGTGCACCTGGATGTGCAAGAACTTTTGGGAGCTTTCAGACAGATCATGATGAGC	445.6
Oy	3908	TGACCACTGACCTTGATGAGCGCGGATTCGTTCTGGAATATGAACTTATCAACCATTC	396.7
Db	4457	TAAACAATGACTTGATGATCGATCAAGAAATCCCTTAACTGCGACTACCGTACCTATGCCATGA	451.6
Oy	3968	GGGTGCTGTTCCAGAAATGAAAGACCACTGTGCTCCGGGACCTTGAGGTCCCGGACT	402.7
Db	4517	GAGTCTGTTCCAGGCAATTGAGAACCACTGTGTTCTGCGGAACTGGAGGTACAGGAA	457.6
Oy	4028	ACCGGCAGAGCGTGTGGAAGAAAGCCTGGAAGCTTTTGCGCCAGCTCATCAACAAGG	408.7
Db	4577	ATGCAAGACAGCACTGGAGAAACCTTGAAACCTTCCGCCAGCTTATCAACAACAAGG	463.6
Oy	4088	TGTTCCCTGCTCCCTTATCCGACAGCGTTGATAGTCCGACGATGCTTCCATAGGCGACCC	414.7
Db	4637	TGTTCTTGCTGACTTATCATCTGACACTGGAACTACACGCGAGCTTCTTCATGCGAAGCC	469.6
Oy	4148	GTGGCAACGTGGCTCACTCATCATGACCGGTGTCGAGACCAAGCTGAGTACGCCACTG	420.7
Db	4697	GTGGGAACGTGGCTCTCTCATATATGACAGGCCCTTTCAGGGTCCGCTAGAAATATCCCACTG	475.6
Oy	4208	ATGTGCTGAGACAGCTGCTGAGCCGACTCTCATTTGACAAAGAACTTGGAGAGCAAGAACCAAC	426.7
Db	4757	ATGTGCTCAAGACAGCTCTCTCTGACCTCATTTGACAAAGAACTTGGAGAAACAAGAACCAAC	481.6
Oy	4268	CTAAGCTGCTGCTCAGAGAGCATGATCACTGTGCTGAGAAAGATCTGACCAATTGTTT	433.7
Db	4817	CCAAGCTGCTTCTCCGAGAGCATGATGCTGTGGCCGAGAAAGATCTGACCTAATCGTTTG	487.6
Oy	4328	CTTTCCTCTCTACAAATCTCTCAAGAGATGTGTGGGAGCCCTCTTCTCCCTGTCTCT	438.7
Db	4877	CTTTTCTTCTACAAAGATCTCTGAAAGATGTGTGGGAGAACCACTCTTCACTGTATATCT	493.6
Oy	4388	GTGTCATCAAGACACATGGAAGAGGCCCATTTGACGCCATCAACGGGACAGGCCGCT	444.7

4937 GTGCATCAAGCAGCAGATG3AAAAAGGCCCATTTGACGCTATTACTGTGTGAGGCCCAT 4996
4448 ACTTCCTTGAGCGAGCAAGCTCATCGCGCAGAGATTGACTACAAAACCTGTGCTGTA 4507
4997 ACTCCCTGAGTGAAGCAAGCTCATCGCGCAGAGATTGACTACAAAACCTGTGCTGTA 5056
4508 GCTGTGTCAAGCCGAGCAATGCCAGAGCCGAGGTCCCAAGTAAATCTTCAACTGTG 4567
5057 ACTGTGTCAAGCCGAGCAATGCCAGAGATCCCAAGTAAATCTTCAACTGTG 5116
4568 ACACCATCATCTAGCTCAGAGAGAGATTTGTGATGCCATCTTCAAGAAATGTCCTGCT 4627
5117 ACACCATCATCTAGCTCAGAGAGAGATTTGTGATGCCATCTTCAAGAAATGTCCTGCT 5176
4628 CCCACCGGCCCAAGCTGAGATATGATCTGAGTGGCGACAGAGAAAGTGGGGCAGAGA 4687
5177 CCCACCGGCCCAAGGGCTGTGAGATCTGAGTGGCGACAGAGAAAGTGGGGCAGAGA 5236
4688 TGATCTTGAGAGATGAAGATATCAACCAAGATTGAGAAATGATTTGAGAGCGACTGAGA 4747
5237 TGGGTGTGAGAGAGAGATTAACCAAAATGAGGTGACTGAGAGCGGCTTAA 5296
4748 CACTGGCCCACTACAGAGTCCAGATGTTCCGTGTGAGATTAAGTGTCCAGAGAGTGA 4807
5297 CACTGATGATTAACAGAGTGTGAGAGATCCGTGTGAGATTAAGTGTCCAGAGAGTGA 5356
4808 CAGCTTAAAGCAGAGATCAACTCCAGCTGTCCAGAGATCTGAGAGATTAAGTGA 4867
5357 CACTCTTAAAGCAGAGATCAACTCCAGCTGTCCAGAGATCTGAGAGATTAAGTGA 5416
4868 ACATGATCCGGGTACAGCGGAGCCGAGAGCTCCGCTCAGAGAGATTAAGTGA 4927
5417 CTTCTCTCAGATTAACAGAGATCCGAGAGATCCGCTCAGAGAGATTAAGTGA 5476
4928 CTGACTGAGAGATGAGATCAAGTGTGAGATCTGAGAGATTAAGTGA 4987
5477 CAGACTTGAAGAGAGATGAGATCAAGTGTGAGATCTGAGAGATTAAGTGA 5536
4988 AGAAGAGAGAGAGATGAGATCAAGTGTGAGATCTGAGAGATTAAGTGA 5047
5537 AGAAGAGAGAGATGAGATCAAGTGTGAGATCTGAGAGATTAAGTGA 5596
5048 CCATTAAGGAGCAGCTGAGAGATTTGTGAGATCTTGTGAGAGATTTGTGAGAG 5107
5597 CCATTAAGGAGCAGCTGAGAGATTTGTGAGATCTTGTGAGAGATTTGTGAGAG 5556
5108 CACACCTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 5167
5657 TGCACCTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 5716
5168 AGGCTGATTAACATGAGATTTGTGAGATCTTGTGAGAGATTTGTGAGAG 5227
5717 AGGAGAGAGAGATTTGTGAGATCTTGTGAGAGATTTGTGAGAG 5776
5228 TGCCCTTGAAGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 5287
5777 TTTCACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 5836
5288 AGAAGAGATTAACATGAGATTTGTGAGATCTTGTGAGAGATTTGTGAGAG 5347
5837 AGGAGAGAGATTTGTGAGATCTTGTGAGAGATTTGTGAGAG 5896
5348 CCACTGTGAGAGATTTGTGAGATCTTGTGAGAGATTTGTGAGAG 5407
5897 CCACTGTGAGAGATTTGTGAGATCTTGTGAGAGATTTGTGAGAG 5956
5408 ACATCCCAAGTGTGAGATTTGTGAGATCTTGTGAGAGATTTGTGAGAG 5467
5957 ATATCCCAAGTGTGAGATTTGTGAGATCTTGTGAGAGATTTGTGAGAG 6016
5468 CCATTAAGGAGCAGCTGAGAGATTTGTGAGATCTTGTGAGAGATTTGTGAGAG 5527
6017 CCATTAAGGAGCAGCTGAGAGATTTGTGAGATCTTGTGAGAGATTTGTGAGAG 6076

5528 AGTTCAACACATGAGTGCATCTCTGAGAGATTTCTTCTTATTTGAGCAATACAGGAG 5587
6077 AGTTCAATATGCTGAGAGGCTCTCAACAGATCTTACTATTTGAGCAATACAGTGAAG 6136
5588 AGATCCTTGAGACCTCTGAGACCAAGATGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 5647
6137 AGCTTATCGGAGACTGAGAGAGATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 6196
5648 TAGAACAAGTATTAACCTCTGAGAGATTTCTTCTTATTTGAGCAATACAGGAG 5691
6197 TGAGCATCTCATCAAGCCATGCTGAGAGATTTCTTCTTATTTGAGCAATACAGGAG 6240

RESULT 2

US-10-087-684-31
Sequence 31, Application US/10087684
Publication No. US20040029116A1

GENERAL INFORMATION:

APPLICANT: Edinger, Shlomit R.
APPLICANT: MacDougall, John R.
APPLICANT: Miller, Isabelle

APPLICANT: Ellerman, Karen
APPLICANT: Stone, David J.

APPLICANT: Grose, William M.
APPLICANT: Lepley, Denise M.

APPLICANT: Rieger, Daniel K.
APPLICANT: Burgess, Catherine E.

APPLICANT: Caeman, Stacie, J.
APPLICANT: Spytek, Kimberly A.

APPLICANT: Boldog, Ferenc L.
APPLICANT: Li, Li

APPLICANT: Padigaru, Muralidhara
APPLICANT: Mishra, Vishnu G.

APPLICANT: Shenoy, Suresh G.
APPLICANT: Rastelli, Luca

APPLICANT: Tchernev, Velizar T.
APPLICANT: Vermet, Corine A.M.

APPLICANT: Zernusen, Bryan D.
APPLICANT: Malpankar, Uriel M.

APPLICANT: Guo, Xiaojia
APPLICANT: Miller, Charles E.

APPLICANT: Gangoli, Esma A.
FILE OF INVENTION: PROTEINS AND NUCLEIC ACIDS ENCODING SAME

FILE REFERENCE: 21402-214 CIP
CURRENT APPLICATION NUMBER: US/10/087,684

CURRENT FILING DATE: 2003-03-10
PRIOR APPLICATION NUMBER: 60/253,834

PRIOR FILING DATE: 2000-11-29
PRIOR APPLICATION NUMBER: 60/250,926

PRIOR FILING DATE: 2000-11-30
PRIOR APPLICATION NUMBER: 60/264,180

PRIOR FILING DATE: 2001-01-25
PRIOR APPLICATION NUMBER: 60/274,194

PRIOR FILING DATE: 2001-03-08
PRIOR APPLICATION NUMBER: 60/313,656

PRIOR FILING DATE: 2001-08-20
PRIOR APPLICATION NUMBER: 60/327,456

PRIOR FILING DATE: 2001-10-05
NUMBER OF SEQ ID NOS: 220

SOFTWARE: CuroSeqList version 0.1
SEQ ID NO 31

LENGTH: 5895
TYPE: DNA

ORGANISM: Homo sapiens
FEATURE:

NAME/KEY: CDS
LOCATION: (23)..(5797)

US-10-087-684-31

Query Match 44.5%; Score 2533.8; DB 12; Length 5895;

Best Local Similarity 68.0%; Pred. No. 0;

Matches 3812; Conservative 0; Mismatches 1712; Indels 81; Gaps 17;

QY 141 CGCGAGGGTTTCATACCTGGTGTGATGAGAGGACAGACACATTACTTGGGGGC 200
Db 217 CGACTGGGGCTCCACCCACTAGTGTGATGAGAGGACAGGAGGTGTATGTGGGCC 276
QY 201 CGTCAATCGATTTTACAAGCTCTCCAGCACTTGAAAGTTTGTGTACGATAGACAG 260
Db 277 AGTGAACCGCATCTATAGCTGTGTGGGAACTGTGACACTGTGCGGGCCACGTCAAG 336
QY 261 GCCGAGAGAGGACAAACCCCAAGTGTACCCACCCGATGTGTACAGACTGTCAATAGCC 320
Db 337 CCTGTGAGAGGACAAACAGAGTGTCTACCCGCCCCAGGTGTCAATCTGTGCCCCACG 396
QY 321 CTTGACCAACCAACAAATGTCAACAAATGTCTCTCATATGACTTACAGAGAAACAGCT 380
Db 397 CTTGGGCAATGTACAAACGTCAACAAAGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 456
QY 381 GATTGCTGTGGGAGCTGTATACCAAGGATCTGTAGAGCTGTGAGGGTGAAGCTCTT 440
Db 457 GCTGGCTGTGGAGCGCTGTCCAGGGCACTGTGCAAGTCTGTGCTGTGAGCAATCTTT 516
QY 441 CAAGCTGGGGAGCTTATCTATAGAAAGAGCACTATCTGTCAAGTGTCAACGAGCGG 500
Db 517 CAATCGGTGTAGCAACCAACCGTAAAGAGCACTACTGTCCAGCGTGCAGAGGACAG 576
QY 501 CTCAGTCTTTGAGTGTGTCTCTTACA-----GCACTGTGATGACAAAGCTGTCA 554
Db 577 CAGCATGGCGGGCTGTCTCATTTGCCGGCCACCGGGCCAGGGCCAAAGCTCTTCTG 636
QY 555 TGCCACGCGAGTGTGAGGAAAGCCGAGTATTTTCCACATCTCCAGCCGGAATGTAC 614
Db 637 GGGACACCCATGATGGCAAGTCCGAGTACTTCCACACTGTCCAGCGGTGGCTCAT 696
QY 615 CAABAATCTTGAGGCGGATGTGACATTTTGGGTAGTCTTCCATGATGAGTTCGTGCTC 674
Db 697 GGCACACGAGAGGATGCCACATGTTGGCTTGTGTACCAAGATAGTTGTGTATC 756
QY 675 GATGATTAAGATCCCTTCGAGCACTTCAACATCATCTGTACTTGTATCTACTATGT 724
Db 757 ACACTCAAGATCCCTTCGAGCACTGTGTCCAAAGTCCCGCCCTTTAACAATCAATGT 826
QY 735 CTATGTTTATGAGTGTGCACTTTGTCTACTTTTATACCTTCAACCTGAGATGTGT 794
Db 817 GTACAGCTTCGCGAGCGAGAGTGTGTCTACTACTACGCTGACGATGACACACAGCT 876
QY 795 TCCACAGGCTCCACACAGAGAGAGGTATATACCAAGCTCGTGAAGCTTTGCA 854
Db 877 GACCTGCGGTATGTGCGCGAGACATTTTCACTGTCAAGATCTGTGGGTCTGTGT 936
QY 855 GAGAGACACAGCTTCACTCTATATGAGGTGCCATTGGCTGTGAGCGCAAGTGGGT 914
Db 937 GACACACCCCAATCTCTCTGTAGTGTATCCCATTTGGCTGACAGAGCGGGGTGT 996
QY 915 GAGATACCGCTGTGAGGCTGTACTGTGTCCAAAGCGGGGCGGTGTGGCAGAGC 974
Db 997 GAGATACCGCTGTGAGGATGTCTTACGCGCGCGCGCGCGGTGTGGCTGTGGCACA 1056
QY 975 CTTTGAATCATCCAGATGATGACCTGTCTTCAACGCTTCTTCAAGGCGCAAGAGCG 1034
Db 1057 GCTGGGCTGTGTGAGAGAGAGAGTGTGTCTGTCACTGTGTGTGTGTGTGTGTGTGT 1116
QY 1035 GAAAAATGAATCCCTGTGATGTGTGCGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1094
Db 1117 CCGCGTGAAGCCACCAAGAGGTCAAGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1176
QY 1095 CCGCATTAAGAGCGGCTGTGAGTCTTGTACCGGGGCGAGGCAAGCTGTGAGCTGTG 1154
Db 1177 GAAATTAAGAGCGGCTGTGAGTCTTGTACCGGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1236
QY 1155 GCTCAAGGTGAAGACATCCCTGTGAGAGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1214
Db 1237 GCTGTCAACAGAGGCTGTGGCTGTGATCACTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1296

QY 1215 TGGCTGGAATGAATGCTCCCTGGGAGTGTCCGACATGTGTGTGTGTGTGTGTGTGT 1274
Db 1297 CGGGCAGGACTTTCAACAGGCTCTGGGGGCGACAGTCAACATTAAGAGGAGCGCTGTGT 1356
QY 1275 CAACGAGGACAGGGACCGCATGACGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1334
Db 1357 GGTGACAAAGATGATGT 1416
QY 1335 GGGCTTTGTGGCAACAAAGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1390
Db 1417 GGTATTCGCGGACAGCAAGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1476
QY 1391 -----GGGCAACGCTCCAGTATGAGAGGTGTGTGTGTGTGTGTGTGTGTGTGT 1445
Db 1477 CGGTGGCGGCTGT 1536
QY 1446 CCGGATATGGCTTCTTCAAGAGACAGAGCAACTTATCATGTGTCAAGAGGAGCT 1505
Db 1537 GCGAGACCTGT 1596
QY 1506 CACCAAGTCTGT 1565
Db 1597 GACGGGT 1656
QY 1566 AGGCAACCCCACTGT 1625
Db 1657 AGGCAACCCCACTGT 1716
QY 1626 TGAAGCTGTCAAGAGAGCGCGGAGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1685
Db 1717 TGAGGAGACAGAGAGCGCGGAGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1776
QY 1686 GGTTCATCCCAATATCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1742
Db 1777 TGTGAGCGCGGAGT 1836
QY 1743 GTACATGTCCCGAGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1802
Db 1837 CTGGAACGT 1896
QY 1803 GGAATGGGCTGT 1862
Db 1897 TGAAGAGT 1956
QY 1863 CCGGATCA-----TCAAGAGATGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1916
Db 1957 GCCCATCAAGCGGGGACAGGGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 2016
QY 1917 AAGAGAGCGGAGAGACCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1976
Db 2017 CAAGAGAGAGGAGAGAGT 2076
QY 1977 CAATTG---TGCCTGT 2033
Db 2077 CAGGTGAGCTGT 2136
QY 2034 GCAATGT 2093
Db 2137 CCAAGT 2196
QY 2094 CGAGAGT 2153
Db 2197 TGAAGAGT 2256
QY 2154 GCTTATCAAGCTGT 2213
Db 2257 ACCCATCAAGCTGT 2316
QY 2214 ATGCATCTTCAACATTCAGGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2273
Db 2317 GTGTCTTGT 2376
QY 2274 CAGGT 2333

Db 2377 CAGCTGCAAGTGCAGAAATTCCTCGATCCTTACAGAGGGAACGATGTACAGGACTGCGC 2436
 QY 2334 CGTGAAGTTGAAGTGTGTGATGAGCACTTCAACATTGAACACCGCTCAAGATTA 2393
 Db 2437 AGTGAACCTGTCACTGTGTGGAACGCACTTTGTCAATTAACAACCAAGAAATCCA 2496
 QY 2394 AGTTCACCTGTAAGATGTGAGGCGATGTGAGAGCTGCGGAGCTGTGCTCAAGCTGA 2453
 Db 2497 GAGGCACTTTAAGATGCGGAGCGCTGTGCGAGAGCTGCGGCTCTGCTTCAAGGCGGA 2556
 QY 2454 CCCAGATTGCAATGTGCTGTGTCAGAGGCGCCAGGCGCAGTGCACCTTGCGCAAGACTG 2513
 Db 2557 CCGCGCTTGAGTGCAGATGTGTGCTGCGCGAGCGCGCTGCTGCTGCGACACCACTG 2616
 QY 2514 CCGTGGCCA---GGAAGCGAGTGGCTGGAGCTGTGTGGCCAAAAGCAATGCAAAA 2570
 Db 2617 CGCTGCGACACACTTGCATCTGTGATGTGACGCGCTCACGGAGAGTGTGCTGACCGA 2676
 QY 2571 CCCCCGATCACAAGATTAATCCCGGTGACAGGCGCCCGGGAAGGGGACCAAGGTAC 2630
 Db 2677 CCCCAGATCTCAAGGTGTGCTGCGAGAGCGGCGCGAGGCGGAGGCGGACCGGCTCAC 2736
 QY 2631 TATCCAGAGGGAAGAACTGTGGCTGTGAATTTGCGGACATGCGCTCCCATGTCAAGGTGC 2690
 Db 2737 TATCAAGAGGGAAGAACTGTGGCTGTGAGTGAAGAGTGTGCTGTGGCGGTGCGCTG 2796
 QY 2691 TGGCGTGAAGTGAAGCTTTAGTGAAGTGTATCATCTCCAGCAAGACAGATGTGTGA 2750
 Db 2797 CAAGTCTGTGAGCGCTTGTGAGAGCGAGTACATCAAGTGTGAGAGCAAGTGTGTGA 2856
 QY 2751 GATGGGGAAGCCCAAGCCC---AGCCAGCATGAGGCTTGTGTGAATCTGTGTGCTGT 2807
 Db 2857 GATGGGGAAGCCCAAGCTGCGTGTGCGCTGATGAGCGCTGTGTGAAGTGTGTGTGGA 2916
 QY 2808 GTGTGCGCTGAATTCATGAGCGCGGTCTCAAGCTTATTAATTGATGACATGACTGT 2867
 Db 2917 CTGTCAACACACTACCGCGCGCTGTGACCCAGCGCTTCACTTGTGTGAACAACCTT 2976
 QY 2868 CTCAGATCTGAAGCCCAAGCGGAGGCCCATGTGTGAGGGAACCAAGTACATCAAG 2927
 Db 2977 CTACCGTGTGAGCCCTCCGCTGTGGCTCTGTGAGGGGGAACCTGATGGATGAGG 3036
 QY 2928 CACCACTGTGAATGCCGGAAGCAAGTGTGTGATGTTTGAAGAGCGCTGTCTT 2987
 Db 3037 AAGCACTGTGAAGAGAGAGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 3096
 QY 2988 CCACAGAGGATCTTCATCTCAATGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 3047
 Db 3097 CTCTGAGTGCAGAGGAATCTCCGTGAGATCCGTGTGTGTGTGTGTGTGTGTGTGT 3156
 QY 3048 GATGAAGT 3098
 Db 3157 TGGGAGCGCTCTCATCATCATCAACATCAACCGGCGGAGTCAACCTGTGAGTGA 3216
 QY 3099 CTTTCAATGTGTGAAGACCCACCATGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 3158
 Db 3217 GTACAACTAACAACGAGAGCCCAACCATCTGAGATGACCCCAAGTGTGAGCATCA 3276
 QY 3159 TGGAAACACACCATTCGCGGTATGGGGGAGCCACCTGAGCTTACAGAAACCCCGAT 3218
 Db 3277 CGGTGGAGCCCTCTGAGGTGTACAGGACCAACCTGTGACCTGTGTGTGTGTGTGT 3336
 QY 3219 CGGTGCAAGATGAGAGGGAAGAGACATCAATATCTGTGAGTGTGTGTGTGTGTGT 3278
 Db 3337 CCGGCGCAAGATGAGAGCATTTAG---AGGAGAACTGTGTGTGTGTGTGTGTGTGT 3393
 QY 3279 GATGACCTGTCAAGGCGCGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 3338
 Db 3394 CATGTATGTGCGCGCGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 3453
 QY 3339 GAGGCCGAGAGT 3398

Db 3454 GCGGCCGATGAGCTGGGCTTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 3513
 QY 3399 GACCACTTCACTTATTCACCAACCGGAGTGTGTGTGTGTGTGTGTGTGTGTGTGT 3458
 Db 3514 CACTCTTCTTCTTACTTACCTTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 3573
 QY 3459 GBACTCAAGCTGTGACCGCTCATCTTAAAGGCGAGAACTGTGTGTGTGTGTGTGT 3518
 Db 3574 GAGCTGAGCCCAAGCTTCCACTGTCTTCAAGGAGCGGAACTTGTGTGTGTGTGT 3630
 QY 3519 TGGGGGCAAGTGAAGCTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGA 3578
 Db 3631 ACCGGCAACTCCGATCTCAACTTACAGGTGTGTGTGTGTGTGTGTGTGTGTGTGT 3690
 QY 3579 CGTGTCAATGT 3638
 Db 3691 CGTGTGAGAGCGCACTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 3750
 QY 3639 GCGCGGT 3698
 Db 3751 GT 3810
 QY 3699 GCTCAGCTGTGCGCATGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 3758
 Db 3811 GCTGAGCTGTGCTCATTTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 3870
 QY 3759 CGTGGCGGT 3818
 Db 3871 GT 3930
 QY 3819 GCTGAGATGTGAGTGAACAACCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 3878
 Db 3931 GCTGAGCTGTGAGTGAACAACCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 3990
 QY 3879 TGGGAGTGTGAGAGCAATCCATGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 3928
 Db 3991 TGCAGAGTGTGAGAGCAATCCATGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 4050
 QY 3929 GTTCTGTGACTATGAACCTTACACCATGTGTGTGTGTGTGTGTGTGTGTGTGTGT 3998
 Db 4051 CTTCTGT 4110
 QY 3999 TGTCTTGT 4058
 Db 4111 TGT 4158
 QY 4059 GCTTGT 4118
 Db 4159 ACTGT 4218
 QY 4119 GTTCCAGGT 4178
 Db 4219 GGCACAGGCGAGCTTGT 4278
 QY 4179 GCTGAGAGCAAGT 4238
 Db 4239 TGAACAAGCTGT 4295
 QY 4339 CGAGAAAGCTGT 4398
 Db 4296 AGT 4355
 QY 4356 GT 4415
 Db 4459 GT 4518
 QY 4416 CCGCATGT 4475
 Db 4519 CCGCATGT 4578

QY 4476 CCAGCAGATTGACTACAAAACCTGGTCTGAGCTGTGTCAAGCCAGCAATGCCAAG 4535
DB 4579 GAGCAGATTGACTACAAAACCTGGTCTGAGCTGTGTCAAGCCAGCAATGCCAAG 4638
QY 4536 CCCGAGGTCCAGTAAGATCCTCACTGTGACACCATCACTCAGGTCAAGAGAGAT 4595
DB 4639 ACTGAGGTGCTGAGAGGGGCTGAGCTGTGACAGGTCACTCAGGCAAGAGAGAT 4698
QY 4596 TCTGAGTGCATCTTCAAGAAATGTGCTGTCTCCACCGGCCCCAAGCTGCAGATATG 4655
DB 4699 GCTGAGCGCTGCTTACAAAGGGCGTGTCTTCTCCAGCGGCCCAAGCGCGAGATG 4758
QY 4656 TCTGAGTGCAGCAAGAGAGTGGGCGAAGATGATCTTGCAGAGATGAAGACATCAC 4715
DB 4759 CTTGAGATGGGCGCAGGGCCGCAATGGGCGCATCTCTGCAAGAGAGAGAGATGCA 4818
QY 4716 CAAGATTGAGATGATTTGAAAGCGACTGAACACACTGGCCCATCTACAGGTGCCAAT 4775
DB 4819 CAAGATTGAGACGATTGGAAGGGCTGAACACACTGGCTCACTACAGGTGACAGACG 4878
QY 4776 TTCCGTGTGGCAATTAATGTCCAGAGGTGACAGGCTTAACGCAAGTGAACAATCCAC 4835
DB 4879 GTCTCGGTGGCAATGTGTGCTCCAGAGAGCTGCTGCTCAACATCTCCATCTCCAC 4938
QY 4836 CGTCTCGAGGACCTCAGCAAGTAATAATGAATCATGATCGGTACACGGGCGAGCCCG 4895
DB 4939 CTTACACCAAG--TCCCTCAGCAGATACGAGACATGTCTGCGACGGCGACGCCCGCA 4995
QY 4896 CAGCTCCGCTCAGGACACCTAATGATCACTCTGACCTGAGAGTGAAGTCAAGATGT 4955
DB 4996 CAGCTCCGCTCAGGACACCTAATGATCACTCTGACCTGAGAGTGAAGTGAAGTGT 5055
QY 4956 GCACCTAGTGAAGAACCAAGCAGCAGGACCAAGAGAGGGGACCGGGGAGCAAGAT 5015
DB 5056 GCACCTAGTGAAGAACCAAGCAGCAGGACCAAGAGAGGGGACCGGGGAGCAAGAT 5115
QY 5016 GGTGTCTGAATCTTACCTGACCCGACTCTGCGCACTA---AGGGGCACTGAGAGAT 5072
DB 5116 GGTGTCTGAATCTTACCTGACCCGACTCTGCGCACTA---AGGGGCACTGAGAGAT 5175
QY 5073 TGTGATGACCTCTTGAAGACATCTTCAAGCAGGCAACCGTGTGCTGCTGCTGCTGCT 5132
DB 5176 TGTGATGACCTCTTGAAGACATCTTCAAGCAGGCAACCGTGTGCTGCTGCTGCTGCT 5235
QY 5133 GGGCATCAAGTACATGTTGACTTCTGAGTGAAGAGGCTGATTAACATGCACTTCA 5192
DB 5236 GGGCATCAAGTACATGTTGACTTCTGAGTGAAGAGGCTGATTAACATGCACTTCA 5295
QY 5193 CCCGAGGTCCAGTAAGATCCTCACTGTGACACCATCACTCAGGTCAAGAGAGAT 5249
DB 5296 TGTGATGACCTCTTGAAGACATCTTCAAGCAGGCAACCGTGTGCTGCTGCTGCTG 5355
QY 5250 CATGATCAAGAACCCGAGTGTGTGATCATCATTAAGAACGATCAAGAGCTG 5309
DB 5356 CATGATCAAGAACCCGAGTGTGTGATCATCATTAAGAACGATCAAGAGCTG 5415
QY 5310 CCTCTGTGTGTGCTGAGACCTTCAAGTCACTTGTGCTGCTGCTGCTGCTGCTGCTG 5369
DB 5416 CCTCTGTGTGTGCTGAGACCTTCAAGTCACTTGTGCTGCTGCTGCTGCTGCTGCTG 5475
QY 5370 CAAGGACTGCTGCTCAACCAAGCTGTGTATGCAAGAGCAATCCCACTTCAAGAAATG 5429
DB 5476 CAAGGACTGCTGCTCAACCAAGCTGTGTATGCAAGAGCAATCCCACTTCAAGAAATG 5535
QY 5430 GGTGA---GAGTATTACTGACATAGGAAAGTGCAGCAGTCAAGAGCAAT 5486
DB 5536 GGTGAAGAGAGTACTTATGACATGCAAGATGCAAGTGCAGCAGTCAAGAGCAAT 5595
QY 5487 GAACGATACCTGCTGAGCAGTCCGAGTGAACATGAATGATTCACACCATGAGTGC 5546
DB 5596 GAACGATACCTGCTGAGCAGTCCGAGTGAACATGAATGATTCACACCATGAGTGC 5655

QY 5547 ACTCTCAGAGATCTTCTCTTATGTGGCAATAACAGCA-----GAGATCTTGGACC 5600
DB 5656 CTTGACAGAGATCTTCTCTTATGTGGCAATAACAGCA-----GAGATCTTGGACC 5715
QY 5601 TCTGACACAGATACCAAGTGTGGGAAGCAAGAACTGGCCATCAAACTAGAACAGTCA 5660
DB 5716 CTTGAGAGAGATAGAGAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 5775
QY 5661 AACCTCATGAGCTTGAACAGCTGA 5685
DB 5776 GAGACAGATGGCCTGAGCAGCTGA 5800

RESULT 3
US-10-218-779-31
Sequence 31, Application US/10218779
Publication No. US20040029222A1
GENERAL INFORMATION:
APPLICANT: Edinger, Shlomit
APPLICANT: MacDougall, John
APPLICANT: Milner, Isabelle
APPLICANT: Ellerman, Karen
APPLICANT: Stone, David
APPLICANT: Gerlach, Valerie
APPLICANT: Grosche, William
APPLICANT: Alshbrook II, John
APPLICANT: Lepley, Denise
APPLICANT: Rieger, Daniel
APPLICANT: Burgess, Catherine
APPLICANT: Casman, Stacie
APPLICANT: Spytek, Kimberly
APPLICANT: Bolog, Ferenc
APPLICANT: Li, Li
APPLICANT: Padigar, Muralidhara
APPLICANT: Mishra, Vishnu
APPLICANT: Patil, Meera
APPLICANT: Shenoy, Suresh
APPLICANT: Rastelli, Luca
APPLICANT: Tcherny, Velizar
APPLICANT: Vernet, Corine
APPLICANT: Zernsen, Bryan
APPLICANT: Malyanar, Uriel
APPLICANT: Guo, Xiaojia
APPLICANT: Miller, Charles
APPLICANT: Gangoli, Esha
TITLE OF INVENTION: Proteins and Nucleic Acids Encoding Same
FILE REFERENCE: 21402-214
CURRENT APPLICATION NUMBER: US/10/218,779
CURRENT FILING DATE: 2002-08-14
PRIOR APPLICATION NUMBER: 60/253,834
PRIOR FILING DATE: 2000-11-29
PRIOR APPLICATION NUMBER: 60/250,-926
PRIOR FILING DATE: 2000-11-30
PRIOR APPLICATION NUMBER: 60/264,180
PRIOR FILING DATE: 2001-01-25
PRIOR APPLICATION NUMBER: 60/313,656
PRIOR FILING DATE: 2001-08-20
PRIOR APPLICATION NUMBER: 60/327,456
PRIOR FILING DATE: 2001-10-05
NUMBER OF SEQ ID NOS: 216
SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO: 31
LENGTH: 5895
TYPE: DNA
ORGANISM: Homo sapiens
US-10-218-779-31

Query Match 44.5%; Score 2533.8; DB 12; Length 5895;
Best Local Similarity 68.0%; Pred. No. 0;
Matches 3812; Conservative 0; Mismatches 1712; Indels 81; Gaps 17;

QY 141 CGCCGAGGTTTCAATCACTGCTGTGTGTGATGAGAGACAGACATTTACTTGGGGG 200

Db	21	GA	CTGGGGG	CGCTCA	CCCA	CCCTAG	GGG	GCAT	GA	GCAG	A	CA	GGG	CGA	GGT	GTAT	GTGG	GGCC	276													
Qy	201	CGT	CAATCG	ATTTT	CA	AGCTT	CT	CA	CCG	CA	CTT	GA	AGT	CTT	GGT	CA	CGA	TGA	CA	GG	260											
Db	277	AGT	GA	CCG	AT	CTAT	AA	GCT	GT	CG	GG	AA	CT	GA	CA	CT	CT	GG	CC	CA	CT	CG	336									
Qy	261	GCC	GA	CGA	GC	AG	CA	CC	CA	AGT	TTT	CC	CA	CC	CG	AT	CGT	CG	CA	GA	CTT	GA	TAG	CC	320							
Db	337	CC	CT	GG	AG	GA	CA	CA	CG	AA	GG	AT	GCT	CA	CC	CG	CC	CAG	GT	CA	GT	CT	CG	CC	CA	CG	396					
Qy	321	CT	TA	CC	CA	CC	CA	CA	AT	GT	CA	CA	GA	AT	GC	T	CC	T	AT	GA	CT	CA	GA	GA	GA	CA	GG	380				
Db	397	CT	CG	GG	CA	GT	A	CT	GA	CA	CGT	CA	CA	AG	CT	G	T	CT	CT	GA	CT	A	T	A	T	GC	CT	AA	CG	456		
Qy	381	GAT	T	G	CC	T	GT	GG	AG	CC	CT	GT	A	CC	AA	GG	C	AT	T	GC	AA	G	CT	GT	GA	AG	CA	CT	TT	440		
Db	457	GCT	GA	CC	T	GT	GG	AG	CC	CG	CT	CC	CA	GG	C	AT	T	GC	AA	TT	C	T	G	T	CT	GT	GC	CA	CT	TT	516	
Qy	441	CA	AG	CT	GG	GG	GA	GC	CTT	A	T	CA	T	GA	AG	GA	GC	A	T	A	T	CT	GT	CA	AG	GT	CA	CA	GA	AG	CG	500
Db	517	CA	AA	CT	GG	GT	AG	CA	CA	CC	CG	TT	AG	GA	GA	CA	CT	A	CT	TC	A	GC	GT	GA	GA	GA	GA	GA	GA	GA	576	
Qy	501	CT	CA	GT	CTT	T	GA	GT	A	T	CGT	CT	CT	CA	-----	GC	AA	CC	T	GA	T	GA	CA	AG	CT	TT	CA	T	554			
Db	577	CAG	AT	GG	CG	GG	CG	GCT	C	ATT	GC	GG	GC	CA	CC	GG	GC	CA	GG	GC	CA	AG	GC	CA	AG	CT	TT	CGT	636			
Qy	555	TG	CC	AC	GG	CA	T	GA	T	GG	AG	CC	CG	AG	AT	TTT	CC	CA	CA	AT	CT	CG	AC	CG	GA	AA	CT	GA	C	614		
Db	637	GG	CA	CA	CC	CA	TC	GA	T	GG	CA	AG	TC	CG	AG	TA	CTT	CC	CA	CA	CT	GT	CA	GC	CG	CT	CA	T	636			
Qy	615	CA	AG	AA	CT	CT	GA	GG	CG	AG	T	GG	C	AT	GT	TC	CG	CA	GT	CTT	CA	T	GA	T	GA	TT	CGT	GC	CT	674		
Db	697	GG	CC	AA	CG	AG	AG	GG	AT	GC	CA	CA	AT	GT	TC	CG	CT	CGT	GA	TC	CA	GA	AT	GA	TT	GT	GT	AT	C	756		
Qy	675	GAT	AT	T	AA	AT	CC	CTT	GG	GA	CA	CC	TT	C	A	CA	CA	T	A	T	CC	CG	A	CTT	GA	T	AT	CT	CA	T	AT	734
Db	757	A	CA	GT	CA	AG	AT	CC	CTT	T	GA	CA	AG	CT	CT	CA	AG	TT	CC	CG	CTT	TT	GA	CA	T	CT	A	T	AT	816		
Qy	735	CT	AT	G	TTT	T	AG	CA	GT	GG	CA	CTT	GT	CT	AT	TTT	GA	CC	T	CA	A	CT	GA	T	AT	GT	GT	C	794			
Db	817	GT	A	C	AG	CTT	CG	CA	CG	AG	CA	GT	T	GT	CT	A	CT	CA	CC	GT	GA	CA	CA	CA	CA	CA	CA	CA	CA	CA	876	
Qy	795	T	CA	CA	CA	GG	CT	CA	CA	CA	GA	GG	CA	GT	GA	T	AT	CA	T	CA	AG	CT	CG	GA	GG	CTT	GA	CA	854			
Db	877	G																														

QY 2334 CGTGGAGTTGACGTCGTGTGGAAATGGGCACTTCAATTGACCAACCCAGCTCAGAAATA 2393
 Db 2437 AGTGAACCTGTCACTGTGTGGAAACGGCACTTTGTCTATGACCAACCCACAGAACTTCA 2496
 QY 2394 AGTTCACTTCTACAACTGTGGACCAATGCGTGAAGCTGCGGCTGTGCTTCAAGCTGA 2453
 Db 2497 GGGCACTCTCTCAAGTGCCTGGCTGCGCGAGACTGCGGCTGTGCTTCAAGCTGA 2556
 QY 2454 CCCAGACTTGGCAATGTGGCTGTGGCCAGAGGCCCCAGGCGCAGTGCACCTGAGGCGACACTG 2513
 Db 2557 CCGCGCTTGAAGTGGATGTGTGCGTGGCCAGCGCGCTGTCTCTCTGACACCACTG 2616
 QY 2514 CCGTGGCCA--GGAGAGCCAGTGTGAGCTGTGTGTGTCMAAAGCAAGTGCACAA 2570
 Db 2617 CGCTGCGGACACACTGTCATCTGTGAATGACAGCGCGCTCAGCGGACAGTGTGCTGACCGA 2676
 QY 2571 CCCCCGATCAAGAAATAATCCCGGTGACAGGCCCCCGGGAAGGGGGGACCAAGTGCAC 2630
 Db 2677 CCCCAGATCTTCAAGCTGTCTCCCGAGAGCGGCCCCGAGCGGGAACGCGCTCAC 2736
 QY 2631 TATCCAGGGGAGAACTTGGGCTGTGAATTTCGCGACATGCGCTCCCATGTCAAGGTTGC 2690
 Db 2737 TATACAGGGGAGAACTTGGGCTGTGGAATTTCGAAAGCTGCGCTGTGGGCGTGGGCTGAG 2796
 QY 2691 TGGGTGTGAGTGCAGCCCTTTAGTGAATGTTTCAATCCCTGACGAAACAGATGTGTGTGA 2750
 Db 2797 CAAGTGTGTGTGAGCCCTCTGTGAAGAGCGATACATCAGTGCAGGACAGATGCTGTGTGA 2856
 QY 2751 GATGGGGGAGGCGCAAGCCC--AGCCAGCATGCAAGGCTTGTGTGAATCTGCGTGTGCTGT 2807
 Db 2857 GATGGGGGAGGCGCAAGCTCTGCGGTGTGCGCATGACGCTGCTGTGAGATGTGTGTGGGA 2916
 QY 2808 GTGTGGCTGTGAATTCTATGAGTGTGCGGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 2867
 Db 2917 CTGTGTACCACTACACGCGCGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 2976
 QY 2868 CTCAATCTGAAGGCGCGGCGGCGGCGCATGTCCGAGGAGGACCCAGATGACATCTACAGG 2927
 Db 2977 CTACGCTGTGAGCTCTCTCCGTTGCGGCTCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 3036
 QY 2928 CACCAACTGTGAATGCCGGAAGCACTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 2987
 Db 3037 AAGCCACTGTGAACCGAGGCAAGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 3096
 QY 2988 CCAAGGCGATCTCAATCTTCAATTTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 3047
 Db 3097 CTCTGT 3156
 QY 3048 GATGAAGGT 3098
 Db 3157 TGGGAGGCTGT 3216
 QY 3099 CTTTCAATGT 3158
 Db 3277 CGGT 3336
 QY 3219 CCGT 3278
 Db 3337 CCGGCGCAAGT 3393
 QY 3279 GATGAAGGT 3338
 Db 3394 CATGT 3453
 QY 3339 GAGGCGCGAGAGTTTGT 3398
 Db 3454 GCGGCGGATGT 3513
 QY 3399 GACCACTTCACTACTATCCCAACCGGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 3458

Db 3514 CACCTCTCTCTTACTATACCTGTGACCCCGTACTGTGAGCCACTGACGCCCTGCTGT 3573
 QY 3459 GGAAGCTCAAGCTGT 3518
 Db 3574 GGAAGCTCAAGCTGT 3630
 QY 3519 TGGGGGCAAGCTGT 3578
 Db 3631 ACCCGCAAGCTGT 3690
 QY 3579 CGTGTCAAGT 3638
 Db 3691 CGT 3750
 QY 3639 GAGCGGT 3698
 Db 3751 GGT 3810
 QY 3699 GCTCAAGCTGT 3758
 Db 3811 GCTCAAGCTGT 3870
 QY 3759 GGT 3818
 Db 3871 GGT 3930
 QY 3819 GCTGT 3878
 Db 3931 GCTGT 3990
 QY 3879 TGGCGAGCTGT 3938
 Db 3991 TGGCGAGCTGT 4050
 QY 3939 GTTCTGT 3998
 Db 4051 GTTCTGT 4110
 QY 3999 TGT 4058
 Db 4111 TGT 4158
 QY 4059 GCTTGT 4118
 Db 4159 ACTGT 4218
 QY 4119 GTTGT 4178
 Db 4219 GGT 4278
 QY 4179 GGT 4238
 Db 4279 CTTGT 4338
 QY 4239 TGAAGAAGCTGT 4295
 Db 4339 CGAGAAGAAGCTGT 4398
 QY 4296 AGT 4355
 Db 4399 GGT 4458
 QY 4416 GGT 4475
 Db 4519 CCGCAATGAGCTGT 4578
 QY 4476 CCGCAATGAGCTGT 4535

Db 4579 GCAGCAGATTGACTACAAGACACTGACCCTGAACTGTGAAACCTTGAAATGAAATGC 4638
Qy 4536 CCCCAGGTCCCAAGTAAGATCTCACTGTGACACCATCACTCAAGGTCAAGAGAAAGAT 4595
Db 4639 AACTAGAGTGGCGGTAAGAGGGCTGAGTGAACGGGTCAACCCAGGCGCAAGAGAACT 4658
Qy 4536 TCTGATGCCATCTTCAAGATATGCTTGTCCACCGGCGCCAAAGCTGCAATATGA 4655
Db 4699 GGTGAGACGTGCTCAAGAGCGGTGCTTACTCCACGCGCCAAAGCGCGGACATGA 4758
Qy 4656 TCTGAGTGGCCACAAGAGAGTGGGCAAGATGATCTTGACAGATGAAGATCAACAC 4715
Db 4759 CTTGAGTGGCCGACAGGGCGCATGCGCATCTCTGACAGAGAGAGCTACAC 4818
Qy 4716 CAAGATTGAGATTTGAGAGAGCTGACACACTGCGCCACTACAGGTGCGAGATG 4775
Db 4819 CAAGATTGACACGATTTGAGAGAGGTGACACACTGCTCACTACAGGTGACAGACG 4878
Qy 4776 TTTCCGTGTGGCATTAGTGTCCAGAGAGTGAACGCTTAAAGCAGTGAACACTCCAC 4835
Db 4879 GTCTTCGTGGGCACTGTGTCCCAAGACAGCTCCGCTTCAACATCTCCACTCTCAC 4938
Qy 4836 CGTCTCCAGACCTGACAGTAATATGAAACATGATCCGGTACACGCGGACGCCGA 4895
Db 4939 CTTACCAAG--TTCCCTACAGAGATACAGAGCATGCTGGCCACAGCCACAGCCCGA 4995
Qy 4896 CAGCCTCCGCTCACGACACTATGATCACTCTGACCTGAGAGTGAATCAAGATG 4955
Db 4996 CAGCCTCCGCTCGCGACCCCATGATCAAGCCCACTGAGAGAGCGGACCAACTGTG 5055
Qy 4956 GACCTTAAGTAAGAACCAAGAGAGGAGACCAAGAGAGAGAGAGAGAGAGAGAGAT 5015
Db 5056 GACCTTAAGTAAGAACCAAGAGAGGAGACCAAGAGAGAGAGAGAGAGAGAGAT 5115
Qy 5016 GGTGTCTGAATCTACCTGACCCGACTCTGCGCACTA--AGGGCACTGACAGAGTT 5072
Db 5116 GGTGTCTGAATCTACCTGACCCGACTCTGCGCACTAAGAGAGAT 5175
Qy 5073 TGTGATGACCTTTTGAAGCATCTTCAAGACAGGACAGGAGTGTGCGCCCTG 5132
Db 5176 TGTGAGCAGCTTTTGAAGCATCTTCAAGACAGGAGAGAGAGAGAGAGAT 5235
Qy 5133 GGCATCAAGTACATGTTGATCTTCTGATGAGAGAGAGAGAGAGAGAT 5192
Db 5236 GGCATCAAGTACATGTTGATCTTCTGATGAGAGAGAGAGAGAGAGAT 5295
Qy 5193 CCGGCACTCCGCACTGAGAGAGAGAT--GCTGCGCTTGAAGTGTGAGTGA 5249
Db 5296 TGTGAGCTGCGCACTGAGAGAGAGAGAGAGAGAGAGAGAGAGAT 5355
Qy 5250 CATGATCAAGACCCGAGATTTGTTGATCTTGAATCAAGAGAGAGAGAGAT 5309
Db 5356 CGTGTATCAAGACCCGAGATTTGTTGATCTTGAATCAAGAGAGAGAGAT 5415
Qy 5310 CTTCTGTGTGAGTCAAGCTTCAATGATCTTCTGCTCAAGTCAAGAGAGAGAT 5369
Db 5416 CTTGTGTGTGAGTCAAGCTTCAATGATCTTCTGCTCAAGTCAAGAGAGAT 5475
Qy 5370 CAAGAGACTGGCCGCAAGAGAGTGTGATGAGAGAGAGAGAGAGAT 5429
Db 5476 CAAGAGACTGGCCGCAAGAGAGTGTGATGAGAGAGAGAGAGAGAT 5535
Qy 5430 GGTGGA--GAGTATTAATCAAGATGAGAGAGAGAGAGAGAT 5486
Db 5536 GGTGAGAGAGAGATTAATCAAGATGAGAGAGAGAGAGAGAT 5595
Qy 5487 GAGAGATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAT 5546
Db 5596 GAGTGTGTGTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAT 5655
Qy 5547 ACTCTCAAGATCTTCTCAATGAGAGAGAGAGAGAGAGAGAT 5600
Db 5656 CTTCAAGAGATCTTCTCAATGAGAGAGAGAGAGAGAGAGAT 5715

Qy 5601 TCTGACACAGATGACCAAGTGTGGAGAGAGAGAGAGAGAGAT 5660
Db 5716 CTTGAGAGAGAGATGACCAAGAGAGAGAGAGAGAGAGAT 5775
Qy 5661 AACCTTCAAGATGAGAGAGAGAGAGAGAGAGAT 5685
Db 5776 GAGACAGATGAGAGAGAGAGAGAGAGAGAT 5800

RESULT 4
US-10-108-260A-802
; Sequence 802, Application US/10108260A
; Publication No. US20040005560A1
; GENERAL INFORMATION:
; APPLICANT: HELIX RESEARCH INSTITUTE
; TITLE OF INVENTION: No. US20040005560A1 full length cDNA
; FILE REFERENCE: H1-A0106
; CURRENT APPLICATION NUMBER: US/10/108,260A
; NUMBER OF SEQ ID NOS: 5458
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 802
; LENGTH: 3666
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-108-260A-802

Query Match 20.3%; Score 1157.8; DB 15; Length 3666;
Best Local Similarity 99.8%; Pred. No. 0;
Matches 1159; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 4527 TGCCAAAGCCCGGAGGTCCAGTAAGATCCTCAACTGAGACCATCACTGAGTCA 4586
Db 1 TGCCAAAGCCCGGAGGTCCAGTAAGATCCTCAACTGAGACCATCACTGAGTCA 60
Qy 4587 GAGAGAGATTTGAGAGAGATCTTCAAGATGCTGCTTCCACCGGCGCCAAAGTGC 4646
Db 61 GAGAGAGATTTGAGAGAGATCTTCAAGATGCTGCTTCCACCGGCGCCAAAGTGC 120
Qy 4647 AGATGATGATGAGAGAGAGAGAGAGAGAGAGAGAGAGATGATCTTGAAGATGAGA 4706
Db 121 AGATGATGATGAGAGAGAGAGAGAGAGAGAGAGAGAGATGATCTTGAAGATGAGA 180
Qy 4707 CATCACCACCAAGATTAAGATTAAGATTAAGATTAAGATTAAGATTAAGATTAAGAT 4766
Db 181 CATCACCACCAAGATTAAGATTAAGATTAAGATTAAGATTAAGATTAAGATTAAGAT 240
Qy 4767 GCCAGATGTTCCGTGTGAGATTAAGATTAAGATTAAGATTAAGATTAAGATTAAGAT 4826
Db 241 GCCAGATGTTCCGTGTGAGATTAAGATTAAGATTAAGATTAAGATTAAGATTAAGAT 300
Qy 4827 CAATCCACGCTCCAGAGACTCAAGCAAGTAAATGAAGATTAAGATTAAGATTAAGAT 4886
Db 301 CAATCCACGCTCCAGAGACTCAAGCAAGTAAATGAAGATTAAGATTAAGATTAAGAT 360
Qy 4887 CAGCCCGGACAGCTCCGCTCAAGAGACCTTAAGATTAAGATTAAGATTAAGATTAAGAT 4946
Db 361 CAGCCCGGACAGCTCCGCTCAAGAGACCTTAAGATTAAGATTAAGATTAAGATTAAGAT 420
Qy 4947 CAAGATGAGCACTTAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAT 5006
Db 421 CAAGATGAGCACTTAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAT 480
Qy 5007 GAGCAAGATGATGATTAAGATTAAGATTAAGATTAAGATTAAGATTAAGATTAAGAT 5066
Db 481 GAGCAAGATGATGATTAAGATTAAGATTAAGATTAAGATTAAGATTAAGATTAAGAT 540
Qy 5067 GAGATTTGTGATGATGATTAAGATTAAGATTAAGATTAAGATTAAGATTAAGATTAAGAT 5126
Db 541 GAGATTTGTGATGATGATTAAGATTAAGATTAAGATTAAGATTAAGATTAAGATTAAGAT 600
Qy 5127 GCGCCGCGCATCAAGATTAAGATTAAGATTAAGATTAAGATTAAGATTAAGATTAAGAT 5186

Db	601	GGCCCTGGCCATCAAGTACATGTTTGAATCTTCTGATATGACAGGCTGTATAACATGGCAT	660
QY	5187	TCATGACCCCGACCTCCGCGCATCTTGGAAGACAAATTGCTGCCCTTGAGGTTTGGCT	5248
Db	661	TCATGACCCCGACCGTCCGCGCATCTTGGAAGACAAATTGCTGCCCTTGAGGTTTGGCT	720
QY	5247	CAACATGATCAAGAACCACCGAGTTGTGTGATCATCATTAAGAACAGATCAACAGACG	5308
Db	721	CAACATGATCAAGAACCACCGAGTTGTGTGATCATCATTAAGAACAGATCAACAGACG	780
QY	5307	CTGCCCTCTCTGTGTGTGCTCAGACCTTCATGACCTCTTGCTTCAGATCAAGACCGGCT	5368
Db	781	CTGCCCTCTCTGTGTGTGCTCAGACCTTCATGACCTCTTGCTTCAGATCAAGACCGGCT	840
QY	5367	GGGCAAGACCTGCGCTCCCAACAAAGCTGTATGCCAAGACATCCCGAGTCAAGAA	5428
Db	841	GGGCAAGACCTGCGCTCCCAACAAAGCTGTATGCCAAGACATCCCGAGTCAAGAA	900
QY	5427	TTGGGTGAGAGGTATTACTCAGACATAGGGAAGATGCCAGCCATCAGGCATCAAGACAT	5488
Db	901	TTGGGTGAGAGGTATTACTCAGACATAGGGAAGATGCCAGCCATCAGGCATCAAGACAT	960
QY	5487	GAACGCATACCTGGCTGAGACAGTCCCGGATGACATGATGATGAGTTCAACACCATGAGTC	5548
Db	961	GAACGCATACCTGGCTGAGACAGTCCCGGATGACATGATGATGAGTTCAACACCATGAGTC	1020
QY	5547	ACTCTCAGAGATCTTCTCTTATGTGGCAAAATACGCGAGAGATCTTTGAACTCTGGA	5608
Db	1021	ACTCTCAGAGATCTTCTCTTATGTGGCAAAATACGCGAGAGATCTTTGAACTCTGGA	1080
QY	5607	CCACATGACCCAGTGTGGGAACGAACAACTGGCTTCAAACTAGAACAACTATACCT	5668
Db	1081	CCACATGACCCAGTGTGGGAACGAACAACTGGCTTCAAACTAGAACAACTATACCT	1140
QY	5667	CATGAGCTTAGCAGCTGAAA	5687
Db	1141	CATGAGCTTAGCAGCTGAGA	1161
RESULT 5			
US-10-245-103-91			
Sequence 91, Application US/10245103			
Publication NO. US20030068778A1			
GENERAL INFORMATION:			
APPLICANT: Baker, Kevin			
APPLICANT: Eaton, Dan			
APPLICANT: Filvaroff, Ellen			
APPLICANT: Goddard, Audrey			
APPLICANT: Grimaldi, J. Christopher			
APPLICANT: Gurney, Austin			
APPLICANT: Smith, Victoria			
APPLICANT: Stephan, Jean-Phillippe			
APPLICANT: Watanabe, Colin			
APPLICANT: Wood, William			
APPLICANT: Zhang, Zemin			
APPLICANT: Fong, Sherman			
TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC			
FILE REFERENCE: P3630R/C112			
CURRENT FILING DATE: US/10/245, 103			
PRIOR FILING DATE: 2002-09-17			
PRIOR APPLICATION NUMBER: 10/197942			
PRIOR FILING DATE: 2002-07-18			
PRIOR APPLICATION NUMBER: 60/059114			
PRIOR FILING DATE: 1997-09-17			
PRIOR APPLICATION NUMBER: 60/063046			
PRIOR FILING DATE: 1997-10-24			
PRIOR APPLICATION NUMBER: 60/065027			
PRIOR FILING DATE: 1997-11-10			
PRIOR APPLICATION NUMBER: 60/079689			
PRIOR FILING DATE: 1998-03-27			
PRIOR APPLICATION NUMBER: 60/086478			

```

: PRIOR FILING DATE: 1998-05-22
: PRIOR APPLICATION NUMBER: 60/087607
: PRIOR FILING DATE: 1998-06-02
: PRIOR APPLICATION NUMBER: 60/089801
: PRIOR FILING DATE: 1998-06-18
: PRIOR APPLICATION NUMBER: 60/090557
: PRIOR FILING DATE: 1998-06-24
: PRIOR APPLICATION NUMBER: 60/090689
: PRIOR FILING DATE: 1998-06-25
: Remaining Prior Application data removed - See File Wrapper or PALM.
NUMBER OF SEQ ID NOS: 116
SEQ ID NO: 91
LENGTH: 2597
TYPE: DNA
ORGANISM: Homo Sapien
US-10-245-103-91

Query Match      13.9%; Score 793.8; DB 14; Length 2597;
Best Local Similarity 99.7%; Pred. No. 1.6e-215;
Matches 795; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 579 CGAGTATTTCCCAACCACTTCACGCGGAAACCTGACCAAGAACTCTGAGCGGTGGAT 638
DB 1 CGAGTATTTCCCAACCACTTCACGCGGAAACCTGACCAAGAACTCTGAGCGGTGGAT 60
QY 639 GTGCGCGGACGCTCCCAATGATGAGTTCGTGGCTCCGATGATTAAGATCCCTTGCGACAC 698
DB 61 GTTGGCGTACCTTTCCATGATGAGTTCGTGGCTCCGATGATTAAGATCCCTTGCGACAC 120
QY 699 CTTCACCAATCACTCCCTGACTTTGATATCTATATGCTATGTTTATGCACTGGCACTT 758
DB 121 CTTCACCAATCACTCCCTGACTTTGATATCTATATGCTATGTTTATGCACTGGCACTT 180
QY 759 TGTCTACTTTTGAACCCCTCCCAACTGAGATGATGATGATGATGATGATGATGATGATGAT 818
DB 181 TGTCTACTTTTGAACCCCTCCCAACTGAGATGATGATGATGATGATGATGATGATGATGAT 240
QY 819 GCAAGGTATATCAATCCAAAGCTCGTGAAGCTTTGCAAGAGAGACCAAGCTTCAACTCTCTA 878
DB 241 GCAAGGTATATCAATCCAAAGCTCGTGAAGCTTTGCAAGAGAGACCAAGCTTCAACTCTCTA 300
QY 879 TGTAGAGGTGCCAATTTGAGCTGTGAGCGCAAGTGGGATGAGTCCGCTGACAGCTGC 938
DB 301 TGTAGAGGTGCCAATTTGAGCTGTGAGCGCAAGTGGGATGAGTCCGCTGACAGCTGC 360
QY 939 CTACTCTGCCAAAGCGGGGCGCGTCTTGGAGAGACCCCTTGGATCCATCCAGTATGATGA 998
DB 361 CTACTCTGCCAAAGCGGGGCGCGTCTTGGAGAGACCCCTTGGATCCATCCAGTATGATGA 420
QY 999 CCTGCTCTTCAACGCTCTTCTTCCAAAGGCGCAGAAACGGAAATGAAATCCCTGATGATGC 1058
DB 421 CCTGCTCTTCAACGCTCTTCTTCCAAAGGCGCAGAAACGGAAATGAAATCCCTGATGATGC 480
QY 1059 GGGCTCTGATCTTTCACTTTGAGGAGATTAAGACCGCATTAAGAGCGGCTGACGTC 1118
DB 481 GGGCTCTGATCTTTCACTTTGAGGAGATTAAGACCGCATTAAGAGCGGCTGACGTC 540
QY 1119 TTGTATACCGGAGGAGGACGCTGGAAGCTGGCCCTGGCTCAAGGTGAAGGACATCCCTG 1178
DB 541 TTGTATACCGGAGGAGGACGCTGGAAGCTGGCCCTGGCTCAAGGTGAAGGACATCCCTG 600
QY 1179 CAGAGAGCGGCTTTAACAATTGACGATTAATTTCTGTGGCTTGACATGAATGTCCTCCCT 1238
DB 601 CAGAGAGCGGCTTTAACAATTGACGATTAATTTCTGTGGCTTGACATGAATGTCCTCCCT 660
QY 1239 GGAAGTGTCCGACATGCTGTGGTGAATTTCCGCTTTCACGAGGACAGAGACCGCATGAC 1298
DB 661 GGAAGTGTCCGACATGCTGTGGTGAATTTCCGCTTTCACGAGGACAGAGACCGCATGAC 720
QY 1299 GTCTGTACATGCAATATGCTTACCAAGAACCACTCTCTGGCTTTTGTGGGACCAAAAGTG 1358
DB 721 GTCTGTACATGCAATATGCTTACCAAGAACCACTCTCTGGCTTTTGTGGGACCAAAAGTG 780

```


PRIOR APPLICATION NUMBER: 60/079689
PRIOR FILING DATE: 1998-03-27
PRIOR APPLICATION NUMBER: 60/086478
PRIOR FILING DATE: 1998-05-22
PRIOR APPLICATION NUMBER: 60/087607
PRIOR FILING DATE: 1998-06-02
PRIOR APPLICATION NUMBER: 60/089801
PRIOR FILING DATE: 1998-06-18
PRIOR APPLICATION NUMBER: 60/090557
PRIOR FILING DATE: 1998-06-24
PRIOR APPLICATION NUMBER: 60/090689
PRIOR FILING DATE: 1998-06-25
Remaining Prior Application data removed - See file wrapper or PALM.
NUMBER OF SEQ ID NOS: 116
SEQ ID NO 91
LENGTH: 2597
TYPE: DNA
ORGANISM: Homo Sapien
US-10-245-143-91

Query Match 13.9%; Score 793.8; DB 14; Length 2597;
Best Local Similarity 99.7%; Pred. No. 1.6e-215;
Matches 795; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 579 CGAGTATTTTCCACCATCTCCAGCCGGAAGTACCAAGAACTCTGAGCGGATGGCAT 638
1 CGAGTATTTTCCACCATCTCCAGCCGGAAGTACCAAGAACTCTGAGCGGATGGCAT 60
QY 639 GTTGGCGTACGCTCTTCATGATGAGTTGTTGGCTCGATGATTAAATCCCTTGGACAC 698
61 GTTGGCGTACGCTCTTCATGATGAGTTGTTGGCTCGATGATTAAATCCCTTGGACAC 120
QY 699 CTTACCATCATCCCTGACTTGTATATCTAATGCTAATGTTTAGAGTGGCACTT 758
121 CTTACCATCATCCCTGACTTGTATATCTAATGCTAATGTTTAGAGTGGCACTT 180
QY 759 TGTCTACTTTTTCAGCCCTCCACCTGATGATGTTCTCCACAGGCTCCACCAAGGA 818
181 TGTCTACTTTTTCAGCCCTCCACCTGATGATGTTCTCCACAGGCTCCACCAAGGA 240
QY 819 GCAGGTATATCATCCAGCTCGTAGAGCTTTTGAAGAGAGACACAGCCTTCACTCTTA 878
241 GCAGGTATATCATCCAGCTCGTAGAGCTTTTGAAGAGAGACACAGCCTTCACTCTTA 300
QY 879 TGTAGAGTGGCCATTGAGTGTGAGCGCAGTGGGTTGAGTACCGCTGCTGACAGCTG 938
301 TGTAGAGTGGCCATTGAGTGTGAGCGCAGTGGGTTGAGTACCGCTGCTGACAGCTG 360
QY 939 CTACCTGTCAGAGCGGGGCGGCTGTTGGCAGAGACCTTGAAGTCCATCAATATGTA 998
361 CTACCTGTCAGAGCGGGGCGGCTGTTGGCAGAGACCTTGAAGTCCATCAATATGTA 420
QY 999 CCTGCTCTTACCGTCTTCTCCAGAGGCGCAAGACGGAATGAATCCCTGGATGATG 1058
421 CCTGCTCTTACCGTCTTCTCCAGAGGCGCAAGACGGAATGAATCCCTGGATGATG 480
QY 1059 GGGCCCTGATCTTCACTTTGAAGCAGATTAATGACCGCATTAAGAGCGGCTGACATC 1118
481 GGGCCCTGATCTTCACTTTGAAGCAGATTAATGACCGCATTAAGAGCGGCTGACATC 540
QY 1119 TTGTTACCGGGGAGAGGAGCGTGAAGCTGGCTCAAGGTGAAGACATCCCTG 1178
541 TTGTTACCGGGGAGAGGAGCGTGAAGCTGGCTCAAGGTGAAGACATCCCTG 600
QY 1179 CAGCAGTGGCTCTTAACCATTAAGATTAATCTTGTGGCTGACATGAATGCTCCCT 1238
601 CAGCAGTGGCTCTTAACCATTAAGATTAATCTTGTGGCTGACATGAATGCTCCCT 660
QY 1239 GGGAGTGGCGACATGAGTGGTGAATTCGGCTTTCACGAGAGACAGGAGCGCATGAC 1298
661 GGGAGTGGCGACATGAGTGGTGAATTCGGCTTTCACGAGAGACAGGAGCGCATGAC 720
QY 1299 GTCTGTATGCAATATGCTTAAGAAACACTCTCTGCTTGTGGGCAACAAAGTGG 1358

DB 721 GTCTGTATGCAATATGCTTAAGAAACACTCTCTGCTTGTGGGCAACAAAGTGG 780
QY 1359 CAGCTGAGAGAGATCC 1375
DB 781 CAGCTGAGAGAGTGC 797

RESULT 8

US-10-245-771-91

Sequence 91, Application US/10245771

Publication No. US20030068781A1

GENERAL INFORMATION:

APPLICANT: Baker, Kevin

APPLICANT: Eaton, Dan

APPLICANT: Filvaroff, Ellen

APPLICANT: Goddard, Audrey

APPLICANT: Grimaldi, J. Christopher

APPLICANT: Gurney, Austin

APPLICANT: Smith, Victoria

APPLICANT: Stephan, Jean-Philippe

APPLICANT: Watanabe, Colin

APPLICANT: Wood, William

APPLICANT: Zhang, Zemin

APPLICANT: Fong, Sherman

TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC

FILE REFERENCE: P1630R1C98

CURRENT APPLICATION NUMBER: US/10/245,771

CURRENT FILING DATE: 2002-09-16

PRIOR APPLICATION NUMBER: 10/197942

PRIOR FILING DATE: 2002-07-18

PRIOR APPLICATION NUMBER: 60/059114

PRIOR FILING DATE: 1997-09-17

PRIOR APPLICATION NUMBER: 60/063046

PRIOR FILING DATE: 1997-10-24

PRIOR APPLICATION NUMBER: 60/065027

PRIOR FILING DATE: 1997-11-10

PRIOR APPLICATION NUMBER: 60/079689

PRIOR FILING DATE: 1998-03-27

PRIOR APPLICATION NUMBER: 60/086478

PRIOR FILING DATE: 1998-05-22

PRIOR APPLICATION NUMBER: 60/087607

PRIOR FILING DATE: 1998-06-02

PRIOR APPLICATION NUMBER: 60/089801

PRIOR FILING DATE: 1998-06-18

PRIOR APPLICATION NUMBER: 60/090557

PRIOR FILING DATE: 1998-06-24

PRIOR APPLICATION NUMBER: 60/090689

PRIOR FILING DATE: 1998-06-25

Remaining Prior Application data removed - See file wrapper or PALM.

NUMBER OF SEQ ID NOS: 116

SEQ ID NO 91

LENGTH: 2597

TYPE: DNA

ORGANISM: Homo Sapien

US-10-245-771-91

Query Match 13.9%; Score 793.8; DB 14; Length 2597;

Best Local Similarity 99.7%; Pred. No. 1.6e-215;

Matches 795; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 579 CGAGTATTTTCCACCATCTCCAGCCGGAAGTACCAAGAACTCTGAGCGGATGGCAT 638
1 CGAGTATTTTCCACCATCTCCAGCCGGAAGTACCAAGAACTCTGAGCGGATGGCAT 60
QY 639 GTTGGCGTACGCTCTTCATGATGAGTTGTTGGCTCGATGATTAAATCCCTTGGACAC 698
61 GTTGGCGTACGCTCTTCATGATGAGTTGTTGGCTCGATGATTAAATCCCTTGGACAC 120
QY 699 CTTACCATCATCCCTGACTTGTATATCTAATGCTAATGTTTAGAGTGGCACTT 758
121 CTTACCATCATCCCTGACTTGTATATCTAATGCTAATGTTTAGAGTGGCACTT 180

PRIOR APPLICATION NUMBER: 60/059114
PRIOR FILING DATE: 1997-09-17
PRIOR APPLICATION NUMBER: 60/063046
PRIOR FILING DATE: 1997-10-24
PRIOR APPLICATION NUMBER: 60/065027
PRIOR FILING DATE: 1997-11-10
PRIOR APPLICATION NUMBER: 60/079689
PRIOR FILING DATE: 1998-03-27
PRIOR APPLICATION NUMBER: 60/086478
PRIOR FILING DATE: 1998-05-22
PRIOR APPLICATION NUMBER: 60/087607
PRIOR FILING DATE: 1998-06-02
PRIOR APPLICATION NUMBER: 60/089801
PRIOR FILING DATE: 1998-06-18
PRIOR APPLICATION NUMBER: 60/090557
PRIOR FILING DATE: 1998-06-24
PRIOR APPLICATION NUMBER: 60/090689
PRIOR FILING DATE: 1998-06-25
PRIOR APPLICATION NUMBER: 60/091358
PRIOR FILING DATE: 1998-07-01
PRIOR APPLICATION NUMBER: 60/091978
PRIOR FILING DATE: 1998-07-07
PRIOR APPLICATION NUMBER: 60/099803
PRIOR FILING DATE: 1998-09-10
PRIOR APPLICATION NUMBER: 60/106932
PRIOR FILING DATE: 1998-11-03
PRIOR APPLICATION NUMBER: 60/115554
PRIOR FILING DATE: 1999-01-12
PRIOR APPLICATION NUMBER: 60/119342
PRIOR FILING DATE: 1999-02-09
PRIOR APPLICATION NUMBER: 60/123957
PRIOR FILING DATE: 1999-03-12
PRIOR APPLICATION NUMBER: 60/123972
PRIOR FILING DATE: 1999-03-11
PRIOR APPLICATION NUMBER: 60/127372
PRIOR FILING DATE: 1999-04-01
PRIOR APPLICATION NUMBER: 60/131271
PRIOR FILING DATE: 1999-04-27
PRIOR APPLICATION NUMBER: 60/133459
PRIOR FILING DATE: 1999-05-11
PRIOR APPLICATION NUMBER: 60/135725
PRIOR FILING DATE: 1999-05-25
PRIOR APPLICATION NUMBER: 60/135729
PRIOR FILING DATE: 1999-05-25
PRIOR APPLICATION NUMBER: 60/135750
PRIOR FILING DATE: 1999-05-25
PRIOR APPLICATION NUMBER: 60/138385
PRIOR FILING DATE: 1999-06-09
PRIOR APPLICATION NUMBER: 60/140653
PRIOR FILING DATE: 1999-06-22
PRIOR APPLICATION NUMBER: 60/141037
PRIOR FILING DATE: 1999-06-23
PRIOR APPLICATION NUMBER: 60/144732
PRIOR FILING DATE: 1999-07-20
PRIOR APPLICATION NUMBER: 60/144758
PRIOR FILING DATE: 1999-07-20
PRIOR APPLICATION NUMBER: 60/144790
PRIOR FILING DATE: 1999-07-20
PRIOR APPLICATION NUMBER: 60/145228
PRIOR FILING DATE: 1999-07-20
PRIOR APPLICATION NUMBER: 60/145698
PRIOR FILING DATE: 1999-07-26
PRIOR APPLICATION NUMBER: 60/146222
PRIOR FILING DATE: 1999-07-28
PRIOR APPLICATION NUMBER: 60/146843
PRIOR FILING DATE: 1999-08-03
PRIOR APPLICATION NUMBER: 60/148188
PRIOR FILING DATE: 1999-08-10
PRIOR APPLICATION NUMBER: 60/148513
PRIOR FILING DATE: 1999-08-12
PRIOR APPLICATION NUMBER: 60/149327
PRIOR FILING DATE: 1999-08-17
PRIOR APPLICATION NUMBER: 60/149395

PRIOR FILING DATE: 1999-08-17
PRIOR APPLICATION NUMBER: 60/150114
PRIOR FILING DATE: 1999-08-20
PRIOR APPLICATION NUMBER: 60/151700
PRIOR FILING DATE: 1999-08-31
PRIOR APPLICATION NUMBER: 60/151734
PRIOR FILING DATE: 1999-08-31
PRIOR APPLICATION NUMBER: 60/162506
PRIOR FILING DATE: 1999-10-29
PRIOR APPLICATION NUMBER: 60/170262
PRIOR FILING DATE: 1999-12-09
PRIOR APPLICATION NUMBER: 60/177118
PRIOR FILING DATE: 2000-01-20
PRIOR APPLICATION NUMBER: 60/179851
PRIOR FILING DATE: 2000-02-02
PRIOR APPLICATION NUMBER: 60/180921
PRIOR FILING DATE: 2000-02-08
PRIOR APPLICATION NUMBER: 60/187202
PRIOR FILING DATE: 2000-03-03
PRIOR APPLICATION NUMBER: 60/198587
PRIOR FILING DATE: 2000-04-18
PRIOR APPLICATION NUMBER: 60/199614
PRIOR FILING DATE: 2000-04-25
PRIOR APPLICATION NUMBER: 60/206330
PRIOR FILING DATE: 2000-05-23
PRIOR APPLICATION NUMBER: 60/206368
PRIOR FILING DATE: 2000-05-23
PRIOR APPLICATION NUMBER: 60/209832
PRIOR FILING DATE: 2000-06-05
PRIOR APPLICATION NUMBER: 60/218371
PRIOR FILING DATE: 2000-07-13
PRIOR APPLICATION NUMBER: 60/222695
PRIOR FILING DATE: 2000-08-02
PRIOR APPLICATION NUMBER: 60/229896
PRIOR FILING DATE: 2000-09-01
PRIOR APPLICATION NUMBER: 60/230621
PRIOR FILING DATE: 2000-09-05
PRIOR APPLICATION NUMBER: 60/232887
PRIOR FILING DATE: 2000-09-15
PRIOR APPLICATION NUMBER: 60/235147
PRIOR FILING DATE: 2000-09-22
PRIOR APPLICATION NUMBER: 60/261878
PRIOR FILING DATE: 2001-01-12
PRIOR APPLICATION NUMBER: 60/261910
PRIOR FILING DATE: 2001-01-16
PRIOR APPLICATION NUMBER: 60/261939
PRIOR FILING DATE: 2001-01-16
PRIOR APPLICATION NUMBER: 60/262150
PRIOR FILING DATE: 2001-01-16
PRIOR APPLICATION NUMBER: 60/264395
PRIOR FILING DATE: 2001-01-25
PRIOR APPLICATION NUMBER: 60/266421
PRIOR FILING DATE: 2001-02-02
PRIOR APPLICATION NUMBER: 60/267623
PRIOR FILING DATE: 2001-02-09
PRIOR APPLICATION NUMBER: 60/274399
PRIOR FILING DATE: 2001-03-09
PRIOR APPLICATION NUMBER: 60/280982
PRIOR FILING DATE: 2001-04-03
PRIOR APPLICATION NUMBER: 60/282129
PRIOR FILING DATE: 2001-04-04
PRIOR APPLICATION NUMBER: 60/282199
PRIOR FILING DATE: 2001-04-04
PRIOR APPLICATION NUMBER: 60/290589
PRIOR FILING DATE: 2001-05-09
PRIOR APPLICATION NUMBER: 60/180997
PRIOR FILING DATE: 1998-11-19
PRIOR APPLICATION NUMBER: 60/267213
PRIOR FILING DATE: 1999-03-12
PRIOR APPLICATION NUMBER: 60/380137
PRIOR FILING DATE: 1999-08-25
PRIOR APPLICATION NUMBER: 60/380138
PRIOR FILING DATE: 1999-08-25

Mon Feb 23 09:34:02 2004

us-09-964-956-12.rmpb

Page 17

PRIOR APPLICATION NUMBER: 09/403297
PRIOR FILING DATE: 1999-10-18
PRIOR APPLICATION NUMBER: 09/423741
PRIOR FILING DATE: 1999-11-10
PRIOR APPLICATION NUMBER: 09/709238
PRIOR FILING DATE: 2000-11-08
PRIOR APPLICATION NUMBER: 09/802706
PRIOR FILING DATE: 2001-03-09
PRIOR APPLICATION NUMBER: 09/872035
PRIOR FILING DATE: 2001-06-01
PRIOR APPLICATION NUMBER: 09/918585
PRIOR FILING DATE: 2001-07-30
PRIOR APPLICATION NUMBER: 09/924419
PRIOR FILING DATE: 2001-08-06
PRIOR APPLICATION NUMBER: 09/927796
PRIOR FILING DATE: 2001-08-09
PRIOR APPLICATION NUMBER: 09/929404
PRIOR FILING DATE: 2001-08-13
PRIOR APPLICATION NUMBER: 09/931836
PRIOR FILING DATE: 2001-08-16
PRIOR APPLICATION NUMBER: 09/941992
PRIOR FILING DATE: 2001-08-28
PRIOR APPLICATION NUMBER: 09/946374
PRIOR FILING DATE: 2001-09-04
PRIOR APPLICATION NUMBER: 10/001054
PRIOR FILING DATE: 2001-11-30
PRIOR APPLICATION NUMBER: 10/052586
PRIOR FILING DATE: 2002-01-15
PRIOR APPLICATION NUMBER: 10/081056
PRIOR FILING DATE: 2002-02-20
PRIOR APPLICATION NUMBER: 10/119480
PRIOR FILING DATE: 2002-04-09

Query Match 13.9%; Score 793.8; DB 14; Length 2597;
Best Local Similarity 99.7%; Pred. No 1.6e-215; Indels 0; Gaps 0;
Matches 795; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 579 CGAGTATTTTCCCAACATCTCCAGCCGGAAGTACCAAGCACTTGAGCGGATGGCAT 638
DB 1 CGAGTATTTTCCCAACATCTCCAGCCGGAAGTACCAAGCACTTGAGCGGATGGCAT 60
QY 639 GTTCGCGTAGCTTTCCATGATGATGTTGGGCTGGATGATTAAGTCCCTCGACAC 698
DB 61 GTTCGCGTAGCTTTCCATGATGATGTTGGGCTGGATGATTAAGTCCCTCGACAC 120
QY 699 CTTACACATCATCCCTGACTTTGATATCTATGTTGATGTTTGAAGTGGCAACTT 758
DB 121 CTTACACATCATCCCTGACTTTGATATCTATGTTGATGTTTGAAGTGGCAACTT 180
QY 759 TGTCTACTTTTGAACCTCTCCAACTGATGATGTTCTCCCAACGGCTCCCAACGAAGA 818
DB 181 TGTCTACTTTTGAACCTCTCCAACTGATGATGTTCTCCCAACGGCTCCCAACGAAGA 240
QY 819 GAGAGTGTATACATCCAAAGCTGAGAGCTTTGCAAGAGGACACAGCTTCAACTCTTA 878
DB 241 GAGAGTGTATACATCCAAAGCTGAGAGCTTTGCAAGAGGACACAGCTTCAACTCTTA 300
QY 879 TGTNAGGTGCGCAATGGCTGTGACGAGTGGGGTGAAGTACCGCTGTGACAGGTGC 938
DB 301 TGTNAGGTGCGCAATGGCTGTGACGAGTGGGGTGAAGTACCGCTGTGACAGGTGC 360
QY 939 CTACTGTGCAAAAGGGGGGCGGCTTGAGAGGACCTTGAGAGTCCATCAATGATGA 998
DB 361 CTACTGTGCAAAAGGGGGGCGGCTTGAGAGGACCTTGAGAGTCCATCAATGATGA 1058
QY 999 CCGTGTCTTCAACCGTCTTCTCCAAAGGCGAGAGCGGAAATGAATCCCTGATGATGTC 1058
DB 421 CCGTGTCTTCAACCGTCTTCTCCAAAGGCGAGAGCGGAAATGAATCCCTGATGATGTC 480
QY 1059 GGGCCGTGCAATCTTCAATGATGATGATGATGATGATGATGATGATGATGATGATG 1118
DB 481 GGGCCGTGCAATCTTCAATGATGATGATGATGATGATGATGATGATGATGATGATG 540

QY 1119 TTGTTACCGGGGGGAGGAGCAAGCTGACCTGAGCTGCAAGTGAAGACATCCCTG 1178
DB 541 TTGTTACCGGGGGGAGGAGCAAGCTGACCTGAGCTGCAAGTGAAGACATCCCTG 600
QY 1179 CAGCAGTGGCTCTTAAACATGACGATTAATCTGCTGAGCATGATGCTCCCT 1238
DB 601 CAGCAGTGGCTCTTAAACATGACGATTAATCTGCTGAGCATGATGCTCCCT 660
QY 1239 GGGAGTTCGACATGGTGGTGAATTCCTGCTTACGAGAGACAGGACGATGAC 1298
DB 661 GGGAGTTCGACATGGTGGTGAATTCCTGCTTACGAGAGACAGGACGATGAC 720
QY 1299 GTCTGTATGCAATGATGATGATGATGATGATGATGATGATGATGATGATG 1358
DB 721 GTCTGTATGCAATGATGATGATGATGATGATGATGATGATGATGATGATG 780
QY 1359 CAAGCTGAAGAGATCC 1375
DB 781 CAAGCTGAAGAGATGC 797

RESULT 12
US-10-238-183-91

Sequence 91, Application US/10238183
Publication No. US20030073189A1

GENERAL INFORMATION:

APPLICANT: Baker, Kevin

APPLICANT: Bacon, Dan

APPLICANT: Filvaroff, Ellen

APPLICANT: Goddard, Audrey

APPLICANT: Grimaldi, J. Christopher

APPLICANT: Gunney, Austin

APPLICANT: Smith, Victoria

APPLICANT: Stephan, Jean-Philippe

APPLICANT: Watanabe, Colin

APPLICANT: Wood, William

APPLICANT: Zhang, Zemin

APPLICANT: Fong, Sherman

TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC

FILE REFERENCE: P3630R1C11

CURRENT APPLICATION NUMBER: US/10/238,183

CURRENT FILING DATE: 2002-09-09

PRIOR APPLICATION NUMBER: 10/197942

PRIOR FILING DATE: 2002-07-18

PRIOR APPLICATION NUMBER: 60/059114

PRIOR FILING DATE: 1997-09-17

PRIOR APPLICATION NUMBER: 60/063046

PRIOR FILING DATE: 1997-10-24

PRIOR APPLICATION NUMBER: 60/065027

PRIOR FILING DATE: 1997-11-10

PRIOR APPLICATION NUMBER: 60/073689

PRIOR FILING DATE: 1998-03-27

PRIOR APPLICATION NUMBER: 60/086478

PRIOR FILING DATE: 1998-05-22

PRIOR APPLICATION NUMBER: 60/087607

PRIOR FILING DATE: 1998-06-02

PRIOR APPLICATION NUMBER: 60/089801

PRIOR FILING DATE: 1998-06-18

PRIOR APPLICATION NUMBER: 60/090557

PRIOR FILING DATE: 1998-06-24

PRIOR APPLICATION NUMBER: 60/090689

PRIOR FILING DATE: 1998-06-25

PRIOR APPLICATION NUMBER: 60/091358

PRIOR FILING DATE: 1998-07-01

PRIOR APPLICATION NUMBER: 60/091978

PRIOR FILING DATE: 1998-07-07

PRIOR APPLICATION NUMBER: 60/099803

PRIOR FILING DATE: 1998-09-10

PRIOR APPLICATION NUMBER: 60/106932

PRIOR FILING DATE: 1998-11-03

PRIOR APPLICATION NUMBER: 60/115554

PRIOR FILING DATE: 1999-01-12

Mon Feb 23 09:34:02 2004

us-09-964-956-12.rnpb

Page 18

PRIOR APPLICATION NUMBER: 60/119342
PRIOR FILING DATE: 1999-02-09
PRIOR APPLICATION NUMBER: 60/123957
PRIOR FILING DATE: 1999-03-12
PRIOR APPLICATION NUMBER: 60/123972
PRIOR FILING DATE: 1999-03-11
PRIOR APPLICATION NUMBER: 60/127372
PRIOR FILING DATE: 1999-04-01
PRIOR APPLICATION NUMBER: 60/131271
PRIOR FILING DATE: 1999-04-27
PRIOR APPLICATION NUMBER: 60/133459
PRIOR FILING DATE: 1999-05-11
PRIOR APPLICATION NUMBER: 60/135725
PRIOR FILING DATE: 1999-05-25
PRIOR APPLICATION NUMBER: 60/135729
PRIOR FILING DATE: 1999-05-25
PRIOR APPLICATION NUMBER: 60/135750
PRIOR FILING DATE: 1999-05-25
PRIOR APPLICATION NUMBER: 60/138385
PRIOR FILING DATE: 1999-06-09
PRIOR APPLICATION NUMBER: 60/140653
PRIOR FILING DATE: 1999-06-22
PRIOR APPLICATION NUMBER: 60/141037
PRIOR FILING DATE: 1999-06-23
PRIOR APPLICATION NUMBER: 60/144732
PRIOR FILING DATE: 1999-07-20
PRIOR APPLICATION NUMBER: 60/144758
PRIOR FILING DATE: 1999-07-20
PRIOR APPLICATION NUMBER: 60/144790
PRIOR FILING DATE: 1999-07-20
PRIOR APPLICATION NUMBER: 60/145228
PRIOR FILING DATE: 1999-07-20
PRIOR APPLICATION NUMBER: 60/145698
PRIOR FILING DATE: 1999-07-26
PRIOR APPLICATION NUMBER: 60/146222
PRIOR FILING DATE: 1999-07-28
PRIOR APPLICATION NUMBER: 60/146843
PRIOR FILING DATE: 1999-08-03
PRIOR APPLICATION NUMBER: 60/148188
PRIOR FILING DATE: 1999-08-10
PRIOR APPLICATION NUMBER: 60/148513
PRIOR FILING DATE: 1999-08-12
PRIOR APPLICATION NUMBER: 60/149327
PRIOR FILING DATE: 1999-08-17
PRIOR APPLICATION NUMBER: 60/149395
PRIOR FILING DATE: 1999-08-17
PRIOR APPLICATION NUMBER: 60/150114
PRIOR FILING DATE: 1999-08-20
PRIOR APPLICATION NUMBER: 60/151700
PRIOR FILING DATE: 1999-08-31
PRIOR APPLICATION NUMBER: 60/151734
PRIOR FILING DATE: 1999-08-31
PRIOR APPLICATION NUMBER: 60/162506
PRIOR FILING DATE: 1999-10-29
PRIOR APPLICATION NUMBER: 60/170262
PRIOR FILING DATE: 1999-12-09
PRIOR APPLICATION NUMBER: 60/171118
PRIOR FILING DATE: 2000-01-20
PRIOR APPLICATION NUMBER: 60/179851
PRIOR FILING DATE: 2000-02-02
PRIOR APPLICATION NUMBER: 60/180921
PRIOR FILING DATE: 2000-02-08
PRIOR APPLICATION NUMBER: 60/187202
PRIOR FILING DATE: 2000-03-03
PRIOR APPLICATION NUMBER: 60/198587
PRIOR FILING DATE: 2000-04-18
PRIOR APPLICATION NUMBER: 60/199614
PRIOR FILING DATE: 2000-04-25
PRIOR APPLICATION NUMBER: 60/206330
PRIOR FILING DATE: 2000-05-23
PRIOR APPLICATION NUMBER: 60/206368
PRIOR FILING DATE: 2000-05-23
PRIOR APPLICATION NUMBER: 60/209832

PRIOR FILING DATE: 2000-06-05
PRIOR APPLICATION NUMBER: 60/218371
PRIOR FILING DATE: 2000-07-13
PRIOR APPLICATION NUMBER: 60/222695
PRIOR FILING DATE: 2000-08-02
PRIOR APPLICATION NUMBER: 60/229896
PRIOR FILING DATE: 2000-09-01
PRIOR APPLICATION NUMBER: 60/230621
PRIOR FILING DATE: 2000-09-05
PRIOR APPLICATION NUMBER: 60/232887
PRIOR FILING DATE: 2000-09-15
PRIOR APPLICATION NUMBER: 60/235147
PRIOR FILING DATE: 2000-09-22
PRIOR APPLICATION NUMBER: 60/261878
PRIOR FILING DATE: 2001-01-12
PRIOR APPLICATION NUMBER: 60/261910
PRIOR FILING DATE: 2001-01-16
PRIOR APPLICATION NUMBER: 60/261939
PRIOR FILING DATE: 2001-01-16
PRIOR APPLICATION NUMBER: 60/262150
PRIOR FILING DATE: 2001-01-16
PRIOR APPLICATION NUMBER: 60/264395
PRIOR FILING DATE: 2001-01-25
PRIOR APPLICATION NUMBER: 60/266421
PRIOR FILING DATE: 2001-02-02
PRIOR APPLICATION NUMBER: 60/267623
PRIOR FILING DATE: 2001-02-09
PRIOR APPLICATION NUMBER: 60/274339
PRIOR FILING DATE: 2001-03-09
PRIOR APPLICATION NUMBER: 60/280982
PRIOR FILING DATE: 2001-04-03
PRIOR APPLICATION NUMBER: 60/282129
PRIOR FILING DATE: 2001-04-04
PRIOR APPLICATION NUMBER: 60/282199
PRIOR FILING DATE: 2001-04-04
PRIOR APPLICATION NUMBER: 60/290589
PRIOR FILING DATE: 2001-05-09
PRIOR APPLICATION NUMBER: 09/180997
PRIOR FILING DATE: 1998-11-19
PRIOR APPLICATION NUMBER: 09/267213
PRIOR FILING DATE: 1999-03-12
PRIOR APPLICATION NUMBER: 09/380137
PRIOR FILING DATE: 1999-08-25
PRIOR APPLICATION NUMBER: 09/380138
PRIOR FILING DATE: 1999-08-25
PRIOR APPLICATION NUMBER: 09/403297
PRIOR FILING DATE: 1999-10-18
PRIOR APPLICATION NUMBER: 09/423741
PRIOR FILING DATE: 1999-11-10
PRIOR APPLICATION NUMBER: 09/709238
PRIOR FILING DATE: 2000-11-08
PRIOR APPLICATION NUMBER: 09/802706
PRIOR FILING DATE: 2001-03-09
PRIOR APPLICATION NUMBER: 09/872035
PRIOR FILING DATE: 2001-06-01
PRIOR APPLICATION NUMBER: 09/918585
PRIOR FILING DATE: 2001-07-30
PRIOR APPLICATION NUMBER: 09/924419
PRIOR FILING DATE: 2001-08-06
PRIOR APPLICATION NUMBER: 09/927796
PRIOR FILING DATE: 2001-08-09
PRIOR APPLICATION NUMBER: 09/929404
PRIOR FILING DATE: 2001-08-13
PRIOR APPLICATION NUMBER: 09/931836
PRIOR FILING DATE: 2001-08-16
PRIOR APPLICATION NUMBER: 09/941592
PRIOR FILING DATE: 2001-08-28
PRIOR APPLICATION NUMBER: 09/946374
PRIOR FILING DATE: 2001-09-04
PRIOR APPLICATION NUMBER: 10/001054
PRIOR FILING DATE: 2001-11-30
PRIOR APPLICATION NUMBER: 10/052586
PRIOR FILING DATE: 2002-01-15

Mon Feb 23 09:34:02 2004

us-09-964-956-12.rnpb

Page 19

PRIOR APPLICATION NUMBER: 10/081056
PRIOR FILING DATE: 2002-02-20
PRIOR APPLICATION NUMBER: 10/119480
PRIOR FILING DATE: 2002-04-09

Query Match 13.9%; Score 793.8; DB 14; Length 2597;
Best Local Similarity 99.7%; Pred. No. 1.6e-215;
Matches 795; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY CGAGATATTTTCCACCATCTCCAGCCGGAACCTGACCAAGAACTCTGAGCGGATGGCAT 638
DB 1 CGAGATATTTTCCACCATCTCCAGCCGGAACCTGACCAAGAACTCTGAGCGGATGGCAT 60
QY GTTCGCGTACGCTCTTCATGATGAGTGTGCGCTCGATGATTAAGATCCCTTGGACAC 698
DB 61 GTTCGCGTACGCTCTTCATGATGAGTGTGCGCTCGATGATTAAGATCCCTTGGACAC 120
QY 699 CTTCACCATCATCCCTGACTTGTATATCTATCTATGTTTGAAGTGGCACTT 758
DB 121 CTTCACCATCATCCCTGACTTGTATATCTATCTATGTTTGAAGTGGCACTT 180
QY 759 TGTCTACTTTTGGACCTTCCAACTGAGATGATGTTCTCCACAGGCTTCCACCAAGAA 818
DB 181 TGTCTACTTTTGGACCTTCCAACTGAGATGATGTTCTCCACAGGCTTCCACCAAGAA 240
QY 819 GCAGGTGTATATCATCCAACTGCTGTGAGCTTTTGCAGAGGACACAGCTTCACTCTTA 878
DB 241 GCAGGTGTATATCATCCAACTGCTGTGAGCTTTTGCAGAGGACACAGCTTCACTCTTA 300
QY 879 TGTAGAGGTGCCCATTTGGCTGTGAGCGAGTGGGAGTACCGCTGCTGAGGCTGC 938
DB 301 TGTAGAGGTGCCCATTTGGCTGTGAGCGAGTGGGAGTACCGCTGCTGAGGCTGC 360
QY 939 CTACCTGTCCAAAGCGGGGGCGGTGCTTGGAGAGACCTTGGAGTCCATCCAGATATGA 998
DB 361 CTACCTGTCCAAAGCGGGGGCGGTGCTTGGAGAGACCTTGGAGTCCATCCAGATATGA 420
QY 999 CCTGCTCTTCAACGCTTCTCCAAAGGGGCGAAGGAGGAAATGAATCCCTGATGATGC 1058
DB 421 CCTGCTCTTCAACGCTTCTCCAAAGGGGCGAAGGAGGAAATGAATCCCTGATGATGC 480
QY 1059 GGCCCTGTGCATCTTTCATCTTGAAGCAGATTAATGACCGCATTAAGAGCGGCTGACATC 1118
DB 481 GGCCCTGTGCATCTTTCATCTTGAAGCAGATTAATGACCGCATTAAGAGCGGCTGACATC 540
QY 1119 TTGTTTACCGGGGGGAGAGGACGCTGAGACCTGAGCTTGCATGAGTGAAGCATCCCTG 1178
DB 541 TTGTTTACCGGGGGGAGAGGACGCTGAGACCTGAGCTTGCATGAGTGAAGCATCCCTG 600
QY 1179 CAGCAGTGGGCTCTTAACCATTAAGATTAATCTTGTGGCTGACATGATGCTCCCT 1238
DB 601 CAGCAGTGGGCTCTTAACCATTAAGATTAATCTTGTGGCTGACATGATGCTCCCT 660
QY 1239 GGGAGTGTCCGACATGCTGCGTGAATTTCCGCTTTCAGAGGAGCAGGAGCCGATGAC 1298
DB 661 GGGAGTGTCCGACATGCTGCGTGAATTTCCGCTTTCAGAGGAGCAGGAGCCGATGAC 720
QY 1299 GTCTGCATCGCATATGTCTACAGAACCACTCTCTGAGCTTTGAGGAGCCAAAGTGG 1358
DB 721 GTCTGCATCGCATATGTCTACAGAACCACTCTCTGAGCTTTGAGGAGCCAAAGTGG 780
QY 1359 CAAAGCTGAAGAGATCC 1375
DB 781 CAAAGCTGAAGAGATGC 797

APPLICANT: Filvaroff, Ellen
APPLICANT: Goddard, Audrey
APPLICANT: Grimaldi, J. Christopher
APPLICANT: Gurney, Austin
APPLICANT: Smith, Victoria
APPLICANT: Stephan, Jean-Philippe
APPLICANT: Watande, Colin
APPLICANT: Wood, William
APPLICANT: Zhang, Zemin
APPLICANT: Fong, Sherman
TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
TITLE OF INVENTION: ACIDS ENCODING THE SAME
FILE REFERENCE: P3630R1C15
CURRENT APPLICATION NUMBER: US/10/238,283
CURRENT FILING DATE: 2002-09-09
PRIOR APPLICATION NUMBER: 10/197942
PRIOR FILING DATE: 2002-07-18
PRIOR APPLICATION NUMBER: 60/059114
PRIOR FILING DATE: 1997-09-17
PRIOR APPLICATION NUMBER: 60/063046
PRIOR FILING DATE: 1997-10-24
PRIOR APPLICATION NUMBER: 60/065027
PRIOR FILING DATE: 1997-11-10
PRIOR APPLICATION NUMBER: 60/079689
PRIOR FILING DATE: 1998-03-27
PRIOR APPLICATION NUMBER: 60/086478
PRIOR FILING DATE: 1998-05-22
PRIOR APPLICATION NUMBER: 60/087607
PRIOR FILING DATE: 1998-06-02
PRIOR APPLICATION NUMBER: 60/089801
PRIOR FILING DATE: 1998-06-18
PRIOR APPLICATION NUMBER: 60/090557
PRIOR FILING DATE: 1998-06-24
PRIOR APPLICATION NUMBER: 60/090689
PRIOR FILING DATE: 1998-06-25
Remaining Prior Application data removed - See File Wrapper or PALM.
NUMBER OF SEQ ID NOS: 116
SEQ ID NO 91
LENGTH: 2597
TYPE: DNA
ORGANISM: Homo Sapien
US-10-238-283-91

Query Match 13.9%; Score 793.8; DB 14; Length 2597;
Best Local Similarity 99.7%; Pred. No. 1.6e-215;
Matches 795; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 579 CGAGATATTTTCCACCATCTCCAGCCGGAACCTGACCAAGAACTCTGAGCGGATGGCAT 638
DB 1 CGAGATATTTTCCACCATCTCCAGCCGGAACCTGACCAAGAACTCTGAGCGGATGGCAT 60
QY 639 GTTCGCGTACGCTCTTCATGATGAGTGTGCGCTCGATGATTAAGATCCCTTGGACAC 698
DB 61 GTTCGCGTACGCTCTTCATGATGAGTGTGCGCTCGATGATTAAGATCCCTTGGACAC 120
QY 699 CTTCACCATCATCCCTGACTTGTATATCTATCTATGTTTGAAGTGGCACTT 758
DB 121 CTTCACCATCATCCCTGACTTGTATATCTATCTATGTTTGAAGTGGCACTT 180
QY 759 TGTCTACTTTTGGACCTTCCAACTGAGATGATGTTCTCCACAGGCTTCCACCAAGAA 818
DB 181 TGTCTACTTTTGGACCTTCCAACTGAGATGATGTTCTCCACAGGCTTCCACCAAGAA 240
QY 819 GCAGGTGTATATCATCCAACTGCTGTGAGCTTTTGCAGAGGACACAGCTTCACTCTTA 878
DB 241 GCAGGTGTATATCATCCAACTGCTGTGAGCTTTTGCAGAGGACACAGCTTCACTCTTA 300
QY 879 TGTAGAGGTGCCCATTTGGCTGTGAGCGAGTGGGAGTACCGCTGCTGAGGCTGC 938
DB 301 TGTAGAGGTGCCCATTTGGCTGTGAGCGAGTGGGAGTACCGCTGCTGAGGCTGC 360
QY 939 CTACCTGTCCAAAGCGGGGGCGGTGCTTGGAGAGACCTTGGAGTCCATCCAGATATGA 998

Db	361	CTACGTGTCAAAGCGGGGGCCGTGCTTGAGGAGACCCTTGAGTGCATCCAGATGATGA	420
Qy	999	CTGCTCTTTCACCGCTCTTCTCCAGAGGCCAGAGCCGAAATTAATCCCTGGATGATGTC	1057
Db	421	CTGCTCTTTCACCGTCTTCTCCAGAGGGCAGAGCGGAAATGAAATCCCTGGATGATGTC	480
Qy	1059	GAGCCGTGATATTCATCTTGAGCAGATGAAATGACCGCATTAAGAGCGAGCTGAGTCC	1118
Db	481	GGCCCTGTGATCTTTATCTTGAGCAGATGAAATGACCGCATTAAGAGCGAGCTGAGTCC	540
Qy	1119	TTGTATCCGGGCGAGGGGACGCTGGACCTGGCTGAGTCAAGGTGAGGACATCCCTCG	1178
Db	541	TTGTATCCGGGCGAGGGGACGCTGGACCTGGCTGAGTCAAGGTGAGGACATCCCTCG	600
Qy	1179	CAGCAGTGGCCTTTACCATTTGAGGATTACTCTGTGTGGCTGGACATGAAATGCTCCCT	1238
Db	601	CAGCAGTGGCCTTTACCATTTGAGGATTACTCTGTGTGGCTGGACATGAAATGCTCCCT	660
Qy	1239	GGAAGTGTCCGACATGATGTCGTGAATTCCTCGTCTTCAGGAGGACAGGACCCGATGAC	1298
Db	661	GGAAGTGTCCGACATGATGTCGTGAATTCCTCGTCTTCAGGAGGACAGGACCCGATGAC	720
Qy	1299	GTCGTATCGCATATGTCATACAGAGACCACTCTCTGGCCTTTGTGGGACCCAAATGTG	1357
Db	721	GTCGTATCGCATATGTCATACAGAGACCACTCTCTGGCCTTTGTGGGACCCAAATGTG	780
Qy	1359	CAAGCTGAAAGATGCC	1375
Db	781	CAAGCTGAAAGATGCC	797

```

RESULT 14
US-10-238-370-91
/ Sequence 91, Application US/10238370
/ Publication No. US2003007319A1
/ GENERAL INFORMATION:
/ APPLICANT: Baker, Kevin
/ APPLICANT: Eaton, Dan
/ APPLICANT: Filvaroff, Ellen
/ APPLICANT: Goddard, Audrey
/ APPLICANT: Grimaldi, J. Christopher
/ APPLICANT: Guiney, Austin
/ APPLICANT: Smith, Victoria
/ APPLICANT: Stephan, Jean-Philippe
/ APPLICANT: Watande, Colin
/ APPLICANT: Wood, William
/ APPLICANT: Zhang, Zemin
/ APPLICANT: Fong, Sherman
/ TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
/ FILE OF INVENTION: ACIDS ENCODING THE SAME
/ FILE REFERENCE: P3630R1C10
/ CURRENT APPLICATION NUMBER: US/10/238,370
/ CURRENT FILING DATE: 2002-09-09
/ PRIOR APPLICATION NUMBER: 10/197942
/ PRIOR FILING DATE: 2002-07-18
/ PRIOR APPLICATION NUMBER: 60/059114
/ PRIOR FILING DATE: 1997-09-17
/ PRIOR APPLICATION NUMBER: 60/063046
/ PRIOR FILING DATE: 1997-10-24
/ PRIOR APPLICATION NUMBER: 60/065027
/ PRIOR FILING DATE: 1997-11-10
/ PRIOR APPLICATION NUMBER: 60/079689
/ PRIOR FILING DATE: 1998-03-27
/ PRIOR APPLICATION NUMBER: 60/086478
/ PRIOR FILING DATE: 1998-05-22
/ PRIOR APPLICATION NUMBER: 60/087607
/ PRIOR FILING DATE: 1998-06-02
/ PRIOR APPLICATION NUMBER: 60/089801
/ PRIOR FILING DATE: 1998-06-18
/ PRIOR APPLICATION NUMBER: 60/090557
/ PRIOR FILING DATE: 1998-06-24
/ PRIOR APPLICATION NUMBER: 60/090689
/ PRIOR FILING DATE: 1998-06-25
/

```

```

? Remaining Seq Application data removed - See File Wrapper or PALM.
?
? NUMBER OF SEQ ID NOS: 116
?
? SEQ ID NO: 91
?
? LENGTH: 2597
?
? TYPE: DNA
?
? ORGANISM: Homo sapien
?
US-10-238-370-91

Query Match      13.9%   SCORE 793.6; DB: 14; Length 2597;
Best Local Similarity      99.7%   Pred. No. 1,6e-215;
Matches 795; Conservative 0; Mismatches 2; Indels 0; Gaps 0

```

QY	579	GGAGTATTTTCCCAACATCTCACCGGAAACTGACCAAGAACTCTGAGCGGAGATGGCAT	638
Db	1	CGAGTATTTTCCCAACATCTCCACCGGAAACTGACCAAGAACTCTGAGCGGAGATGGCAT	60
QY	639	GTTCCGCGTAGCTCTTCATGATGATGATTCGTGCGCTCGATGATTAAGATCCCTTCGGACAC	698
Db	61	GTTCGCGAGCATCTTCATGATGATGATTCGTGCGCTCGATGATTAAGATCCCTTCGGACAC	120
QY	699	CTTCACCACATACCCCTGACTTTGATPACTACTATGTCATAGTATTACAGTGGCACTT	758
Db	121	CTTCACCACATACCCCTGACTTTGATPACTACTATGTCATAGTATTACAGTGGCACTT	180
QY	759	TGTCCTACTTTTGAACCTTCCAACTGAGATGTTGTCTCCACAGGCTTCACCAACCAAGAA	818
Db	181	TGTCCTACTTTTGAACCTTCCAACTGAGATGTTGTCTCCACAGGCTTCACCAACCAAGAA	240
QY	819	GCAGGTGATPACATCCAAAGCTGCTGAGAGCTTTGCAAGAGAGGACACAGGCTTCACATCCTTA	878
Db	241	GCAAGGTGATPACATCCAAAGCTGCTGAGAGCTTTGCAAGAGAGGACACAGGCTTCACATCCTTA	300
QY	879	TGTAGAGGTGCCATTTGCTGTGTGACGCGAGTGGGAGTGAACCGCTGCTACAGGCTGC	938
Db	301	TGTAGAGGTGCCATTTGCTGTGTGACGCGAGTGGGAGTGAACCGCTGCTACAGGCTGC	360
QY	939	CTACTGCTCCAAAGCGGAGGCGCTGTGTGACAGACCTTTGAGATCCATCCAGATGATGA	998
Db	361	CTACTGCTCCAAAGCGGAGGCGCTGTGTGACAGACCTTTGAGATCCATCCAGATGATGA	420
QY	999	CCTGCTCTTCAACGCTCTCTCCAAAGGCGCAAGCGGAAATGAATCCCTGATGAGTTC	1058
Db	421	CCTGCTCTTCAACGCTCTCTCCAAAGGCGCAAGCGGAAATGAATCCCTGATGAGTTC	480
QY	1059	GAGCCCTGTGCATCTTCACTTGTGAAGCAGATTAATGACCCGATTAAGAGAGCCGCTGCAGTTC	1118
Db	481	GAGCCCTGTGCATCTTCACTTGTGAAGCAGATTAATGACCCGATTAAGAGAGCCGCTGCAGTTC	540
QY	1119	TTGTATCCGGGGCGAGGAGCAGCGTGGACCTGAGCCGTGCTCAAGGTGAAGAGCATCCCGTG	1178
Db	541	TTGTATCCGGGGCGAGGAGCAGCGTGGACCTGAGCCGTGCTCAAGGTGAAGAGCATCCCGTG	600
QY	1179	CAGCAGTGCAGTCTTAAACATTGACGATTAACCTTGTGTGCTGACATGATGCTCCCT	1238
Db	601	CAGCAGTGCAGTCTTAAACATTGACGATTAACCTTGTGTGCTGACATGATGCTCCCT	660
QY	1239	GGGAGTGTCCGACATGATGATGCGTGGAAATCCCGTCTTCACGGAGGACAGGAGCCGATGAC	1298
Db	661	GGGAGTGTCCGACATGATGATGCGTGGAAATCCCGTCTTCACGGAGGACAGGAGCCGATGAC	720
QY	1299	GTCGTCTCATGCGCATATGTCATAAAGAACCATCTCTGTGGCCTTTGTGTGGGACCAAAAGTGG	1358
Db	721	GTCGTCTCATGCGCATATGTCATAAAGAACCATCTCTGTGGCCTTTGTGTGGGACCAAAAGTGG	780
QY	1359	CAAGGTGAAGAAAGATCC 1375	
Db	781	CAAGGTGAAGAAAGTGC 797	

Mon Feb 23 09:34:02 2004

us-09-964-956-12.inpb

Page 21

GENERAL INFORMATION:
APPLICANT: Baker, Kevin
APPLICANT: Eaton, Dan
APPLICANT: Filvaroff, Ellen
APPLICANT: Goddard, Audrey
APPLICANT: Grimaldi, J. Christopher
APPLICANT: Guirney, Austin
APPLICANT: Smith, Victoria
APPLICANT: Stephan, Jean-Philippe
APPLICANT: Watanabe, Colin
APPLICANT: Wood, William
APPLICANT: Zhang, Zemin
APPLICANT: Feng, Sherman
TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
FILE REFERENCE: P3630R1C88
CURRENT APPLICATION NUMBER: US/10/245,055
CURRENT FILING DATE: 2002-09-16
PRIOR APPLICATION NUMBER: 10/197942
PRIOR FILING DATE: 2002-07-18
PRIOR APPLICATION NUMBER: 60/059114
PRIOR FILING DATE: 1997-09-17
PRIOR APPLICATION NUMBER: 60/063046
PRIOR FILING DATE: 1997-10-24
PRIOR APPLICATION NUMBER: 60/065027
PRIOR FILING DATE: 1997-11-10
PRIOR APPLICATION NUMBER: 60/079689
PRIOR FILING DATE: 1998-03-27
PRIOR APPLICATION NUMBER: 60/086478
PRIOR FILING DATE: 1998-05-22
PRIOR APPLICATION NUMBER: 60/087607
PRIOR FILING DATE: 1998-06-02
PRIOR APPLICATION NUMBER: 60/089801
PRIOR FILING DATE: 1998-06-18
PRIOR APPLICATION NUMBER: 60/090557
PRIOR FILING DATE: 1998-06-24
PRIOR APPLICATION NUMBER: 60/090689
PRIOR FILING DATE: 1998-06-25
Remaining Prior Application data removed - See File Wrapper or PALM.
NUMBER OF SEQ ID NOS: 116
SEQ ID NO 91
LENGTH: 2597
TYPE: DNA
ORGANISM: Homo Sapien
US-10-245-055-91

Query Match 13.9%; Score 793.8; DB 14; Length 2597;
Best Local Similarity 99.7%; Pred. No. 1.6e-215;
Matches 795; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 579 CGAGTATTTTCCACCATCTCCAGCCGGAACCTGACCAAGAACTCTGAGGCGAGTGCAT 638
DB 1 CGAGTATTTTCCACCATCTCCAGCCGGAACCTGACCAAGAACTCTGAGGCGAGTGCAT 60
QY 639 GTTCGGGTACGTTTCCATGATGAGTTCGTGGCTTCGATGAAGATCCCTTGGAGAC 698
DB 61 GTTCGGGTACGTTTCCATGATGAGTTCGTGGCTTCGATGAAGATCCCTTGGAGAC 120
QY 699 CTTACCATCATCCCTGACTTGTATCTACTATGTCATGTTTGTAGCAGTGGCACTT 758
DB 121 CTTACCATCATCCCTGACTTGTATCTACTATGTCATGTTTGTAGCAGTGGCACTT 180
QY 759 TGTCTACTTTTGAACCTTCCAACTGAGATGTTCTCCACGAGTCCACCAAGGA 818
DB 181 TGTCTACTTTTGAACCTTCCAACTGAGATGTTCTCCACGAGTCCACCAAGGA 240
QY 819 GCAGGTGTATACATCCAGGCTGTGAGGCTTTGCAAGAGGACACAGCCTTCACTCTTA 878
DB 241 GCAGGTGTATACATCCAGGCTGTGAGGCTTTGCAAGAGGACACAGCCTTCACTCTTA 300
QY 879 TGTAGAGGTGCCATTGGCTGTGAGGCGAGTGGGGTGGAGTACCGCTTCTGAGGCTGC 938
DB 301 TGTAGAGGTGCCATTGGCTGTGAGGCGAGTGGGGTGGAGTACCGCTTCTGAGGCTGC 360

QY 939 CTACCTGTCCAAAGCCGGGGGCGTGTGTCAGAGACCTTTGAGATCCATCCAGATATGA 998
DB 361 CTACCTGTCCAAAGCCGGGGGCGTGTGTCAGAGACCTTTGAGATCCATCCAGATATGA 420
QY 999 CCGTCTTTTACCGTCTTCTCCAAAGGCGCAAGAGCGGAAATGAAATCCCTGGATGATC 1058
DB 421 CCGTCTTTTACCGTCTTCTCCAAAGGCGCAAGAGCGGAAATGAAATCCCTGGATGATC 480
QY 1059 GGCCTGTGCATCTTATCTTGAAGAGATTAATGACCCCATTTAGAGAGCGGTGATGTC 1118
DB 481 GGCCTGTGCATCTTATCTTGAAGAGATTAATGACCCCATTTAGAGAGCGGTGATGTC 540
QY 1119 TTGTTACCGGGGAGAGGACGCTGGAAGCTGGCTGCTCAAGGTGAAGACATCCCTG 1178
DB 541 TTGTTACCGGGGAGAGGACGCTGGAAGCTGGCTGCTCAAGGTGAAGACATCCCTG 600
QY 1179 CAGCAGTGCCTCTTAAACATTAAGCATTAATTCTGTGGCTTGACATGAATGCTCCCT 1238
DB 601 CAGCAGTGCCTCTTAAACATTAAGCATTAATTCTGTGGCTTGACATGAATGCTCCCT 660
QY 1239 GGGAGTGTCCGATGATGTTGTTGGAATTCCTCCCTTCAAGGAGAGAGGACCGCATGAC 1298
DB 661 GGGAGTGTCCGATGATGTTGTTGGAATTCCTCCCTTCAAGGAGAGAGGACCGCATGAC 720
QY 1299 GTCTGTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1358
DB 721 GTCTGTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 780
QY 1359 CAAGCTGAAGAAATTC 1375
DB 781 CAAGCTGAAGAAATTC 797

Search completed: February 20, 2004, 02:47:35
Job time : 1250 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: February 19, 2004, 16:15:27 ; Search time 8735 Seconds

(without alignments)
19455.715 Million cell updates/sec

Title: US-09-964-956-12

Perfect score: 5691
Sequence: 1 atgaagccatgccctggaa.....gcttagacagctgaataaa 5691

Scoring table:

IDENTITY NUC
Gapop 10.0 , Gapext 1.0

Searched: 27513289 seqs, 14931090276 residues

Total number of hits satisfying chosen parameters: 55026578

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

EST:*
1: em_estsdb:*
2: em_estsdb:*
3: em_estsdb:*
4: em_estsdb:*
5: em_estsdb:*
6: em_estsdb:*
7: em_estsdb:*
8: em_estsdb:*
9: em_estsdb:*
10: em_estsdb:*
11: em_estsdb:*
12: em_estsdb:*
13: em_estsdb:*
14: em_estsdb:*
15: em_estsdb:*
16: em_estsdb:*
17: em_estsdb:*
18: em_estsdb:*
19: em_estsdb:*
20: em_estsdb:*
21: em_estsdb:*
22: em_estsdb:*
23: em_estsdb:*
24: em_estsdb:*
25: em_estsdb:*
26: em_estsdb:*
27: em_estsdb:*
28: em_estsdb:*
29: em_estsdb:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	2239.6	39.4	5691	29	AY420219 Homo sapi
2	2175.4	38.2	3361	11	AK051614 Mus muscu
3	2111.4	37.1	5564	29	AY420221 Mus muscu
4	2017.2	35.4	5520	29	AY420666 Homo sapi

Result No.	Score	Query Match	Length	DB ID	Description
5	1767.2	31.1	5447	29	AY420668 Mus muscu
6	1681.4	29.5	5520	29	AY420667 Pan trogl
7	1552.6	27.3	5341	29	AY420220 Pan trogl
8	1221.2	21.5	3355	11	AK049319 Mus muscu
9	743.8	13.1	838	14	CD653539 AGENCOURT
10	658.8	11.6	1059	12	BM547417 AGENCOURT
11	650.6	11.4	1201	9	AL529477 AL529477
12	638.4	11.2	745	14	CA316867
13	613	10.8	618	13	EX470595
14	599.2	10.5	1027	12	BM544169
15	578	10.2	882	13	BU149880
16	556.4	9.8	646	14	CD352102
17	550.4	9.7	862	14	CB724079
18	546.4	9.6	924	10	BF312056
19	538.8	9.5	638	12	BC066250
20	525.2	9.2	831	14	CD349170
21	525.2	9.2	932	13	BU121882
22	517.6	9.1	822	14	CD802878
23	514.2	9.0	821	14	CD802734
24	513.8	9.0	867	13	EX723674
25	512.8	9.0	1037	13	BU117252
26	510.8	9.0	762	13	BU268254
27	505.4	8.9	612	10	BB656076
28	502.4	8.8	760	14	CF533779
29	501.4	8.8	844	13	BO434703
30	500.2	8.8	798	14	CA752328
31	499.8	8.8	733	14	CD351785
32	498.8	8.8	864	13	EX720011
33	498.6	8.8	657	13	EX670344
34	495.4	8.7	762	9	AL134739
35	494.6	8.7	653	9	BE294546
36	493.8	8.7	740	14	CA328914
37	493.8	8.7	976	13	BU250116
38	492.8	8.7	725	12	BM950962
39	491.6	8.6	711	14	CB519051
40	491.6	8.6	838	12	BT824913
41	488.4	8.6	716	14	CB249549
42	487.4	8.6	789	13	EX320194
43	486.4	8.5	702	12	BG825770
44	485.8	8.5	817	14	CB525253
45	484.6	8.5	894	14	CA488958

ALIGNMENTS

RESULT 1
AY420219
LOCUS
DEFINITION Homo sapiens HMW151 gene, VIRTUAL TRANSCRIPT, partial sequence.
ACCESSION AY420219
VERSION AY420219.1 GI:39776176
KEYWORDS GSS.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
REFERENCE 1 (bases 1 to 5691)
Clark,A.G., Glanowski,S., Nielson,R., Thomas,P., Kejariwal,A., Todd,M.A., Tanenbaum,D.M., Civeille,D.R., Lu,F., Murphy,B., Ferreira,S., Wang,G., Zheng,X.H., White,T.J., Shinsky,J.J., Adams,M.D. and Cargill,M.
Inferring nonneutral evolution from human-chimp-mouse orthologous gene trios
JOURNAL Science 302 (5652), 1960-1963 (2003)
PUBMED 14671302
REFERENCE 2 (bases 1 to 5691)
Clark,A.G., Glanowski,S., Nielson,R., Thomas,P., Kejariwal,A., Todd,M.A., Tanenbaum,D.M., Civeille,D.R., Lu,F., Murphy,B., Ferreira,S., Wang,G., Zheng,X.H., White,T.J., Shinsky,J.J., Adams,M.D. and Cargill,M.
Direct Submission

Mon Feb 23 09:34:03 2004

us-09-964-956-12.rst

Page 2

JOURNAL Submitted (16-NOV-2003) Celera Genomics, 45 West Gude Drive,
Rockville, MD 20850, USA
COMMENT This sequence as made by sequencing genomic exons and ordering them
based on alignment.
FEATURES
Location/Qualifiers
1..5691
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
<1..5691
/locus_tag="HOM7151"
ORIGIN
Query Match 39.4%; Score 2239.6; DB 29; Length 5691;
Best Local Similarity 58.0%; Pred. No. 0;
Matches 3228; Conservative 0; Mismatches 2299; Indels 42; Gaps 8;
QY 141 CGCGAGGGTTTCATCATCCTGGTGGTGGATGAGAGACAGACATTTACTTGGGGC 200
DB 141 CGACTGGGGCCTACCCACTTAAGTGGTGAAGAGACAGAGAGTGTATGTGGGGC 200
QY 201 CGTCAATCGATTATCAAGCTCTCCAGCGACTGAAGTCTTGGTGAAGCATGAGACAG 260
DB 201 AGTGAACCGCATCTATAGCTGTGGGGAACTGACACTGTGGGGCCACGTACCGG 260
QY 261 GCGGAGAGAGACACCCCAAGTTTACCACCCCGCATGTCCAGACCTGCATATGAGCC 320
DB 261 CCTGTGAGAGACACAGAAAGTGTCTACCGCGCCAGGTGCACTTGGCCCCAGG 320
QY 321 CCTGACCAACCAACAATGTCAACAAGATGCTCTCATAGACTACAGAGAAACAGCT 380
DB 321 CTGGGAGTACTGACAAAGCTCAACAAGCTGTGCTGAGACTATGCGCTAACCGCT 380
QY 381 GATTGCTTGGGAGCCTGTACCAAGGCATCTGCAAGCTGAGGCTGAGAGACCTT 440
DB 381 GCTGGCTTGGGAGCGCCTCCAGGCGATCTGCAAGTCTGCTGCTGAGCAATTTCT 440
QY 441 CAAGCTGGGGAGCCTTATCATAGAGAGACACTATCTGTCAAGTGTCAACGAGACGG 500
DB 441 CAACCTGGGTGAGCCACACCAACCGTAAAGAGCACTACTGTCAAGTGTCAAGAGCAG 500
QY 501 CTGAGCTTTGGAGTATCTGTCTTCTAC-----AGAACCTGGAAGCAAGTGTCT 554
DB 501 CAGCATGGCGGCGCTGTCAATTCGGGCGACCGGGCAAGGCAAGCTCTTGGT 560
QY 555 TGCCACGGCAGTGGATGGAGAGCCCGAGTATTTTCCACCATCTCCAGCCGGAACCTGAC 614
DB 555 GGGGACACCCATCATGAGCAAGTCCGAGTACTTCCCACTGACCGCGTGGCTCAT 620
QY 615 CAAGAACTCTGAGCGGATGAGCATGTTGGCGTACGTTCTTCAAGTAAAGTCTGGCCCTC 674
DB 615 GGGCAACAGAGAGAGATCCGACATGTTGGCTTGTGTACAGAGTGAAGTTGTGCATC 680
QY 675 GATGATTAAGATCCCTTCGAGCACTTCAACCATCATCCCTGACTTTGATATCAATATG 734
DB 675 ACAGCTCAAGATCCCTTCGAGCAAGTCTGCTCAAGTTCGGCTTTGATATCAATATG 740
QY 735 CTATGTTTAAAGAGTGGCAACTTTGTCTACTTTTGAACCTTCAACCTGAGATGAGTGT 794
DB 735 GTACAGCTTCGAGCGACAGAGATTTGTCTACTACTACGCTGAGTAAACACACAGCT 800
QY 795 TCCACCAAGGCTCCACCAAGAGAGAGTGTATATCAACAGCTGTGAGCTTTGCAA 854
DB 801 GACTCGCTGATGCGCGCGCGAGCACTTCTTCAAGTCCAAATGCTGGGCTGTGTGT 860
QY 855 GAGAGACACAGCCTTCAACTCTATGTAGAGTGCCTATGGCTGTGAGCGCAGTGGGCT 914
DB 861 GAGAGACCCCAAAATTTCTACTGTACTGTAGTTCCCATTTGGCTGCGAGAGCGGGTGT 920
QY 915 GGAATTAACGCTGAGCAGGCTGCTACTGCTTCCAAAGCGGGGCGCTGCTTGGCAGAG 974
DB 921 GGAATTAACGCTGAGCAGGCTGCTACTGAGCGCGCGCGCTGCTTGGCAGAG 980

QY 975 CTTTGAAGTCAATCAATGATGACCTGCTCTTCAACCGTCTTCTCAAGGGCCAGAGCG 1034
DB 981 GCTGGGCGCTGAGTGAAGAGAGAGCGTGTCTTCACTGTGTGCGCCAGGGCCAGAGAA 1040
QY 1035 GAAATAGAAATCCCTGATAGTGGCGCTGTGATCTTCACTTGAAGAGATTAATGA 1094
DB 1041 CCGGTGAACCAACAAAGAGTGAAGTCACTGTCCCTGTACAGCTCAGGGCCATCAAGA 1100
QY 1095 CCGCATTAAGAGCGCTGAGTCTTGTATCCGGGGCGAGAGGACAGCTGGAACCTGGCCG 1154
DB 1101 GAAATTAAGAGAGGATCAAGTCTGCTACCGTGTGAGAGGAGCAAGCTCTTCCCTG 1160
QY 1155 GCTCAAGTGAAGGACATCCCTGAGCAAGTGGCGCTTTAACCATGACGATTAATCTTG 1214
DB 1161 GCTGCTCAACAGGAGTGGGCTGTGATCACTGANNNNNNNNNNNNNNNNNNNNNNNN 1220
QY 1215 TGCGCTGACATGATGCTCCCTGGAGTGTCCGACATGTCGTGAATTCGCCCTT 1274
DB 1221 NNN 1280
QY 1275 CACGAGAGACAGGACCGCATGACGTCTGTATGTGATATGTCTACAGAACCACTCTCT 1334
DB 1281 NNN 1340
QY 1335 GGCCTTTGTGGGACCAAAAGTGGCAAGCTGAAGAGATCCGGGTGATGACCA--- 1390
DB 1341 NNN 1400
QY 1391 -----GGGGAACGCCCTCCAGTATGAGCGGTGCAAGTGGTGAACCCGGGCCAGTCT 1445
DB 1401 CGTGGCGCGCTCCCTGCGCTACGAGAGGCTGTGAGCCAGAGAGGACAGCCCATCT 1460
QY 1446 CCGGATATGCGCTTCTCAAGGACCAAGCACTTACATCATGTACAGAGAGCAGCT 1505
DB 1461 GCGAGACCTGCTCTCAGGCCCAACCAACAGTACTCTTACGCCATGACAGAGAGCAGT 1520
QY 1506 CACCAAGTCCCTGTGAGTCTGTGTGATATCAAGCTGGCGGAGAGTCTGTGCTC 1565
DB 1521 GACGGGTGTCTGTGAGAGCTGTGTGACATCAAGTCTGTGAGCTGTGTGCTGGCTC 1580
QY 1566 AGCGACCCCACTGTGAGTGTGTGTGCTGACAAACGTGCAACCCGGAAGAGCGGTG 1625
DB 1581 ACGGACCCCACTGTGAGTGTGTGTGCTGACAAAGTCTGTGAGCTGTGTGCTGGCTC 1640
QY 1626 TGAGCGGTCAAGAGAGCCCGCAGTGTGCTGCGAATGAAGAGTGTGCGGCTGAC 1685
DB 1641 TGAGCGAGACAGAGAGAGCCCGCAGTGTGCTGCGAATGAGTGTGAGCTGAC 1700
QY 1686 GGTCCATCCCAACATATCTGCTCTCAGTAAACAGTGTGTGCTGTGAGAGCTA 1745
DB 1701 TGTGAGCCCGCATGTGTGTGTGACATGTGTGACAGTGTGCTGAGAGCTG 1760
QY 1746 CAATGTCCGAGCTGTGAGTGTGCTGCACTTGAAGAGCTGTGAGAGAGTGA 1805
DB 1761 GAACGTGCTCACTGTGAGTGTGCTGCACTTGTGAGAGAGCTTCAAGATGTGA 1820
QY 1806 TGGCGTGTGTGAGGCAATGATTCAGATGAGTGTGCTGCGTGCAGAGAGTGTGCGCC 1865
DB 1821 GAGCTGCTGAGAGATGAGCGGATTCAGTGTGCTGCGTGCAGAGAGTGTGCGCC 1880
QY 1866 GATCATCAAGAGATGAGAGAGCAACATGTGTGACAGTGTGCAATTAAGAGAGAC 1925
DB 1881 CATCAAGCGGAG 1940
QY 1926 CCGCAGTACCTTGGCAGACACAGCTTGTCTTCAATTAAGAGAGAGAGAGAGAGAG 1985
DB 1941 AAGGAGAGAGT 2000
QY 1986 CCGT 2045
DB 2001 CCGT 2060
QY 2046 CCATGACCCCAAGACCTGCTCTTCAAGAGAGCGAGTGAAGTGTGCGGAGAGAGCTGCC 2105

D	2061	ACACAAAGTGGCTGACTGGCGCTTCCCTGGAGGGCCGATGTCAAGTGTCTGAGGACTGCC	2120
Q	2106	CCAGCTCTCTCGAGTGGACAAAGATCTGTGTGCCGTGAGAGTGTACAAGCTATCAAGCT	2165
D	2121	ACAGATCTCTGGCTTCCACGAGAACTTACGTGTGAGTGGAGTGTAAACCTATCAAGCT	2180
Q	2166	GAAAGCCAAAGAACTCCCGCCAGCCAGCTGTGGACACGTAGCTACAAATGCAATCCCAA	2225
D	2181	GGCGCCACGGAACCTTGCCACAGCCACAGTCAAGCTACGTGTGATATGATGTCCCTTCCA	2240
Q	2226	CATTCAAGGACGCGAGACGAGTCCCGCTTGCCTTCAACAGCTCCAGCGTCAAGT	2285
D	2241	CATCCCGGGACGCCGGCCGTGTCAACGCCCTCGCTTCAACAGCTCCAGCGTCAAGT	2300
Q	2286	CCAAACAACCTCTTATTCTATGAAGGATGGAATCAACAACCTGCCCGTGAAGTTGAC	2345
D	2301	CCAAATTTCTCGNN	2360
Q	2346	AGTGTGTGAATGGCACTTCAACATTGACAAACCAGCTCAGATATAAGTTCACTCTA	2405
D	2361	NN	2420
Q	2406	CAAGTGTGAGCCATGCGTGAAGCTGGGGCTGTGCTCAAGGCTACCCAGACTTCCG	2465
D	2421	NN	2480
Q	2466	ATGTGTGTGTGTCAGAGGCCACAGGCAGTGACCTGTCCGACACTGCCCTGCCACGA	2525
D	2481	NN	2540
Q	2526	GAGCAGTGTGTGAGCTGTCTGTGCCAAAAGCAAGTGACAAACCCCGCATACAGA	2585
D	2541	NN	2600
Q	2586	GATATA---CCCGGTGACAGGCCCCCGGGGAAGGGGGACCAAGGTCACTATCCAGAGGA	2642
D	2601	NNNNCTGTCCCCGAGAGGGGCCCGAGGCAAGGGGCGACGGGCTCACTATCAAGGCGA	2660
Q	2643	GAACTTGGGCTGGAATTTGCGCAACATGCGCTCCCATGTCAAGGTTGCTGGCGTGAAGT	2702
D	2661	GAACTGGGCTGTGGATTGGAAGAGTGTGCTGTGAGCGTGGCGTGGCAAGGTGTGTG	2720
Q	2703	CAGGCTTTAAGTGAATGTTATTCCTTCGACGAACAATCGTGTGTGAAGATGGGGAGGC	2762
D	2721	CAGGCTGTGAGAGCGAGTCAATCAATGTCGGAACAAGTGTGTGTGAAGATGGGGAGGC	2780
Q	2763	CAAGGCC---AGCCAGCAATGACAGCTTGTGTGAATCTGCTGTGTGTGTGTCGGCTGA	2819
D	2781	CAGCTCCGTGTGTGCCATGACGCCCTGTGTGAAGTGTGTGTGGGGAGTGTCTCAACCA	2840
Q	2820	ATTCAATGCGCCGTCTTCAAGCTCTTATTTACTTATGACACTGACTCTCTCAATCTGAA	2879
D	2841	CTACCGCGCCCTGTACCCCAAGCGCTTCACTTGTGTGACAAACCTTCTACCGTGTAG	2900
Q	2880	GCCCAACCGGGGGCCCATGTCCGAGGGAGCCCAAGTGACAAATCAAGGACCAACTGAA	2939
D	2901	CCCTTCCGTGGGCTCTGTGAAGGGGACACTGTGATTTGACATCAAGGAGACCACTGAA	2960
Q	2940	TGCCGGAAGCAACGTGTGTGATATGTTTGAAGAGCAAGCTGTCTTTTCCACAGGCGATC	2999
D	2961	CGAAGGAGTATGT	3020
Q	3000	TTCATCTCAATGTGTCCAAACCACTCTCAATAGAGTGTGATGAAGATGAAGTGTCTC	3059
D	3021	CTCCGTGAGATCCGATGCTGTGACACCCCGGGGACAGCCCTGACAGCCTCCCATAT	3080
Q	3060	GGTGAGGTGACAGAGGCCAAG---ATCCACACAGAACTGTGCTTTCAGTATGTGAAGA	3116
D	3081	CATCAACATCAACGGCGCCAGGCTCACCAACCTGAGGTGAAGTCAACTCTACACCGAGA	3140
Q	3117	CCCCCACTGTGTGCGGATTTGAAGCCAGAAATGAGCACTTGTCAATGGAACAACCAATGC	3176

D	3141	CCCCCACCATCCTGGATGATGACCCCCGAGTGGAGCATCAACGNNNNNNNNNNNNNNNNNNNN	3200
Q	3236	CGTATGGGGGAGCCACCTGGAGCTCATACAGAAACCCCAATCCGTGCCAAGATGGAGG	3236
D	3201	NN	3260
Q	3237	GAAAGACACATCATATATCTGTGAGGTTCTGAAACGCTACTAGATGACCTGTGAGCGCC	3296
D	3261	NN	3320
Q	3297	CGCCCTGGCTCTGGGTCCTGAACCAAGTCAGACTGACCCGAGAGGCCCCGAGAGATTGG	3356
D	3321	GTCTGTGGCCAAACCCCTGTGGCAACCCACCAAGACTGGGAGAGCGCGGATAGCTGGG	3380
Q	3357	CTTATCCTGGAGCAAGCTCAGTCCCTGCTCATCTCTCAAGACCAACTTCACTTCA	3416
D	3381	CTTGATGATGACACCTGGGCTCCCTGCTTGTGCTCAATCCACCTCTTCCCTTACTA	3440
Q	3417	TCCCAACCCGGTGTGTAAGGCTTTTGGTCCCTCAGAACTCTGAGACTCAAGCTGGAC	3476
D	3441	CCCTGACCCCGTACTGAGGCCACTCACCCCACTGGCGCTCTGAGAGTGAAGCCACACTC	3500
Q	3477	GCCCATATCCTTAAGGGCAAGAACCGATCCCGCCTGTGGCTGGGGGCAACGTGAAGCT	3536
D	3501	CCCACTACTCTCAAGGACCGGAACCTCTTGCCACT--GCAACCGGCAACTCCGACT	3557
Q	3537	GAACTACACTGTCTGGTTGGGGAAGACCGGTGACCCGTGACCGTGTCAATGTCCAGCT	3596
D	3558	CAACTACACGGTGTCTATCGGCTCCACACCTGTGACCTGACCGTGTGCGAGACCACT	3617
Q	3577	GCTCTGGAGTCCCCCAACTCATTCGGCAGGCAAAAGTATGGCCGCTGCGGTGGCAT	3656
D	3618	GCTGTGGAGCGGCCCACTCATCTGTGGACGCAAAAGTCAAGTGTGGGCGAGTGGCTT	3677
Q	3657	GGAATTACTCCCGGGAGATGGTATCATTTGCCCCGACAGCCGCTGCAAGCTGCCCGCAT	3716
D	3678	GGAATTCCTCGCAAGGACATGCAAGGTATCTGGACAGCCTGTGACGCTGCTGCAT	3737
Q	3717	CGTCAGCATCGCAATGGCTGGCGGCTTCCTCATATTTTCATCGTGGCCGTCATTTGC	3776
D	3738	TGTGGGCAATGGCGAGCGGGGGTCTCTGCTGCTGATCTGTGCTGTGCTAATGCG	3797
Q	3777	CTATTAACGCAAGTCCCGCGAAATGACCTCACCGTGAAGCGGCTGAGATGACATGGA	3836
D	3798	CTTACAGGCAAGTCAAGAAATGCTGACCGACACTCAAGCGCTGACGCTCCAGATGGA	3857
Q	3837	CAACTGAGTCCCGTGTGGCCCTGAGTGCAGAGTCAAGAAACCTTTGCCGAGCTGCACGGA	3896
D	3858	CAACTGGAGTCCCGCGTGGCCCTCGAATGCAAGAAACCTTTGCAAGCTGCAGACAGA	3917
Q	3897	CATCATATAGTGAACAAGTGAACCTGATGAGCCGGGATTCGTTCTGCACTATAGAC	3956
D	3918	CATTCACAGCTCAATGACTGTGAGCGGTGCCGATCCCTCTCTGACTACACCGAC	3977
Q	3957	TTAACAATCGGGGTGCTGTTCCAGAGATTGAAGCAACCTGTCTCTCCGGACCTTGA	4016
D	3978	ATATGCCATGCGGGGTCTCTTCTGGGATGAGAGCAACCTGTGCTCAAGAGATGGA	4037
Q	4017	GGTCCCGGACTACCGGAGAGAGCGGTGAGAAAGCCTGGAAGCTTTGGCCAGCTCAT	4076
D	4038	GGT-----GAGGGCAATTGGAAGATGCTGACACTGTTGGGAGCACTGCT	4085
Q	4077	CAACAACAAGGTGTTCTGCTCTCTTATTCGACAGCTTGAGTCCAGCGTAACTTCTC	4136
D	4086	GACCAAGAAGCACTTCCTGTGACTTTCATTCGACAGCTGAGGACACAGCCAGCTTCTC	4145
Q	4137	CATGGGCAACGCTGGCAACGTGGCTCACTATCATGACGCTGCGAGCAAGCAAGCTGA	4196
D	4146	CATGGGCAACGCGGGAATGTGGCTGCTCATATGACGGCCCTGCAAGGCGAGATGA	4205
Q	4197	GTAACCACTATGTGCTGAAGAGAGCTGTGCGCACTCATTTGACAAAGAACTTGAAG	4256
D	4206	ATAGGCCACAAGCGGTGCTCAAGAGCTGCTTTCCGACTCATGCAAGAAACTTGAAG	4265

425	CAGAACACACCCCTAAGCTGCTGCTCAGAGAGACTGATCAGTGGCTAGAAAGTGTGAC	4316
426	CAGAACCAACCCCAAGCTGTCTACTGGCCCGGACTGATCTGGTGGCAGAAAGATGTCAAC	4325
4317	CAATTGGTTACTTTCCTCTCTACAAAGTTCTCAAGAGTGTGCTGGGAGCCCTCTT	4376
4326	TAACTGGTTACCTTCTCTGTATTAAGTTCTTCAAGAAATGGCTGGGAGCGCGTGT	4385
4377	CTCCCTGTCTGTSCCATCAAGCAGCAGATGAGAAAGGCGCCATTGACCCCATCAAGG	4436
4386	CATCTGTACTGTGCCCATCAAGCAGCAGATGAGAAAGGCGCCATTGACCCCATCAAGG	4445
4437	CGAGGCCGCTACTTCCTTGAAGCGAGCAAGCTCATCCGCGACAGATTGACTTCAAAAC	4496
4446	TGAGGCAACGCTACTCCCTGATGTAGGCAAGCTCATCCGCGACAGATTGACTTCAAGAC	4505
4497	CCTGTCCTGAGCTGTGTCAAGCCGACCAATGCGCAAGCGCCGAGGTCCAGTAAAGAT	4556
4506	ACTGACCTGTAACTGTGTGAACCTGTAGAAATGAGAAATGACCTGAGGTGGCGTAAAGG	4565
4557	CTTCACTGTGACCACTACTCAAGTGTGAAGAGAAATTTGGATGCCATCTTCAAGAA	4616
4566	GCTGAGACTGTGACACGGTCAACCGGCGCAAGAGAACTGTGTGACCTGCTTCAAGAGG	4625
4617	TGTGCTTTGTCTCCACCGGCGCCAAAGCTGCAGATATGATCTGAGTGTGCGACAGAGAG	4676
4626	CGTGCCCTAANNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNAAGTGGCGGACGGCGG	4685
4677	TGGGGCAAGATGATCTTTGCAGATGAAGACTACCAACCAAGTTGAGATGATTGGAA	4736
4686	CATGGCGGCGACTACTCTCGCAGAGCAGAGGAGTCAACACCAAGATTGACAACGATTGAAA	4745
4737	GCGACTGAACACACTGGGCCACTACAGAGTCCAGATGAGTTCCGTGGTGCAATTAAGTGC	4796
4746	GAGGTGAACACTGTGCTCACTCAAGGTACAGAGGGGTCTCGGTGGGCACTGTGTGC	4805
4797	CAAGCAGTGAACAGCTTATTAACGAGTGAACAATCCACCGTCTTCAGAGCTTACAGAG	4856
4806	CAAGCAGACGTCCGCTTCAACAATCTCCAACTCTTCACTTCAACAAG--TCCCTCAG	4862
4857	TAAATATAAAAACATGATCCCGGTACACGGGACGCGCCGACAGCTCCGCTCAACGAGACCC	4916
4863	CAGATACGAGAGCACTGTGCGCAAGCGCCAGACCGCCGACAGCTCGGCGTGGCGACGCC	4922
4917	TATGATCACTCTGCACTTGAAGTGAAGTGAATCAAGATGTGCGACCTTATGAGAAACACGA	4976
4923	CATGATCACGCCCGACTGTGAAGAGGAGACCAAGAGTGTGACACCTGTGAAAGAACACGA	4982
4977	GCAACGAGACAGAAAGAGAGGAGGAGACCGGAGGAGCAAGATGTGTGAATCTTACCTGAC	5036
4983	CCACCTGAGACAGCGGTGAAGGTGACCGCGGACGAAGATGTGTGAGATCTTACTTGAC	5042
5037	CCGATCTCTGCGCACTTAAGGGGACACTGTCAAAATTTGTGATGATACCTCTTTGAGACAT	5096
5043	ACGGGTACTGTGGCAACAAAGGGGACACTGCAAAATTTGTGACGACCTGTTGAGACAT	5102
5097	CTTCAACACGAGCACCGTGTGCTCGCCCTGCGCCCTGCGCATCAAGTACATGTTGACTT	5156
5103	CTTCAACACGAGCACCGGGGTCAAGCCCTGCGCTGCGCATCAAGTACATGTTGACTT	5162
5157	CTTGAATGACAGGCTGTATTAACATGCAATTCATGATCCCGCAACGTCGGCCATCACTGGAA	5216
5163	CTTGATGAGACGGCGGACAAACACCAAGTCAAGATCTGACGTGGCGCAACCTTGGAA	5222
5217	GAGCAATTCGCTGGCCCTGAGGTTTGGGTCAACATGATCAAGAACCCGACGTTTGTGT	5276
5223	GAGCAATTCGCTGGCCCTGAGGTTTGGGTGAACTGATCAAGAACCCACGTTTGTGT	5282
5277	TGACATCCATAGAACAGATACAGAGCGCCGCTCTGTGAGTGGCTCAGACCTTAT	5336
5283	CGACATTCACAAAGACAGATACGAGCGCTGTTGTGGTGTGGCGCCACGACTTAT	5342

[illegible]

TITLE
JOURNAL
REFERENCE
AUTHORS

Analysis of the mouse transcriptome based on functional annotation of 60,770 full-length cDNAs
Nature 420, 563-573 (2002)
(bases 1 to 3361)
6 (bases 1 to 3361)
Aach, J., Alzawa, K., Akimura, T., Arakawa, T., Bono, H., Carninci, P., Fukuda, S., Furuno, M., Hanagaki, T., Hara, A., Hashizume, W., Hayashida, K., Hayatsu, N., Hiramoto, K., Hiraoka, T., Hirozane, T., Hori, F., Imotani, K., Ishii, Y., Itoh, M., Kagawa, I., Kasukawa, T., Kato, H., Kawai, J., Kojima, Y., Kondo, S., Kono, H., Kouda, M., Koyama, S., Kurihara, C., Matsuyama, T., Miyazaki, A., Murata, M., Nakamura, M., Nishii, K., Nomura, K., Numata, R., Ohno, M., Ohse, N., Okazaki, Y., Saito, R., Saitoh, H., Sakai, C., Sakai, K., Sakazume, N., Sano, H., Sasaki, D., Shibata, K., Shinagawa, A., Shiraki, T., Sogabe, Y., Tagami, M., Tagawa, A., Takahashi, F., Takaku-Akahira, S., Takeda, Y., Tanaka, T., Tomaru, A., Toya, T., Yasunishi, A., Muramatsu, M. and Hayashizaki, Y.

FEATURES
SOURCE

Location/Qualifiers

1..3361
/organism="Mus musculus"
/mol_type="mRNA"
/strain="CS7BL/6J"
/db_xref="PANTOM,DB:DJ30060F15"
/db_xref="WGI:2419923"
/db_xref="taxon:10090"
/clone="DJ30060F15"
/tissue_type="spinal ganglion"
/clone_lib="Riken full-length enriched mouse cDNA library"
/dev_stage="12 days embryo"
1..2559
/note="unlabeled protein product; putative
similar to PLEXIN 2 [Mus musculus] (SPTR/P70207, evidence:
FAST, 78.4%ID, 45.1%length, match=2553)"
/codon_start=1
/protein_id="BAC34692.1"
/db_xref="GI:26342060"
/translation="VRIPEWMSISGNTPIAVWGTHTDILIONPIRAKGGKEHINIC
EVLNATERTCPALALGPHQSDLTERRPERGILLINUNSLINNTTYNPNP
EAFSPGILELPGPIILKGNLIPPAAGNVAANTVAVGEPCVTYTSVDLCE
SPNLGRKVAAGVGEYSPGNVIAVDSPLPALVASAGGLITLIVALLN
KRKRESDLTKRLQMDNLESVALECKEFAELDTDLHELSDLDGAGIPLDR
TYKRVLPGLIEDHVLADLEPGVROEVRKGLPAOLINNVKPLSLRTESVA
ESMDRGNVASLINTVOSLEKAVTDLADLIDKNLNSKPLSLRTESVA
EKLNTNTPILLYKEKAGEPLSLFCALKKOKEGPDIAIGEARYSLSDEKLR
COLEKTVLSCVSPDNVNSPEVPIKINCTTOVKEKTLDAIKVNPSSHPKRA
NDLEWQSGARMILQDEDTTKLENBKRKLNTAHYOVDPGSVALVSKQVRYAA
NNSVTSRTSARKYENMLRYTSPDSLSRTWITPDSGVKHLVKNHSHGQSG
DRGKRWSEIYLRLATKTKLQKFEVDLFEFTISTARHSALPLAKVQFDECA
DKGHIDPHVHTKSNCLPLRFVNMINKPQFVDIHKNSITDACISAAQFMDSC
STSERLKQDSPNKLVAYDIPIKMYVERVYDIDGMPAISIDQNMNAAIABSRNH
MNEFYMALSEISYVYKSEIIGPLDHDQCKOKLAKYKEQVITLMSLS"

ORIGIN

Query Match 38.2% Score 2175.4; DB 11; Length 3361;
Best Local Similarity 90.6%; Pred. No. 0;
Matches 2320; Conservative 0; Mismatches 241; Indels 0; Gaps 0;

3127 GTGGGATTTGACCGAATGAGCATTTGTCAAGTGAACAACACCCATCGCGTATGGGG 3186

1 GTGGGATTTGACCGAATGAGCATTTGTCAAGTGAACAACACCCATCGCGTATGGGG 60
3187 ACCGACCTGACCTCATACAGAACCCCGAGTCCGTGCAAGCATGAGGAGAGAC 3246
61 ATTCACCTGACCTCATACAGAACCCCGAGTCCGTGCAAGCATGAGGAGAGAC 120
3247 ATCAATATCTGAGAGTTCTGAAAGCTTACTGAGATGACCTGTGAGCGCCCTGCT 3306
121 ATCAATCTGAGAGTTCTGAAATCTTACAGAGATGACCTGTGAGCGCCCTGCT 180
3307 CTGGGTCTGACCACTGACAGCTGACAGAGAGAGAGAGAGAGAGAGAGAGAGAG 3366
181 CTGGGTCTGACCACTGACAGCTGACAGAGAGAGAGAGAGAGAGAGAGAGAGAG 240
3367 GACACCTGACAGCTGACAGCTGACAGAGAGAGAGAGAGAGAGAGAGAGAGAG 3426
241 GACACCTGACAGCTGACAGCTGACAGAGAGAGAGAGAGAGAGAGAGAGAGAG 300
3427 GTGTTTGAAGCTTTGTTGCTTCTGAGAACTCTGAGAGCTGAGAGAGAGAGAG 3486
301 GTGTTTGAAGCTTTGTTGCTTCTGAGAACTCTGAGAGCTGAGAGAGAGAGAG 360
3487 CTAAAGGAG 3546
361 CTAAAGGAG 420
3547 GTGCTGTTGAG 3606
421 GTGCTGTTGAG 480
3607 TCCCCCACTGATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 3666
481 TCCCCCACTGATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 540
3667 CCGGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 3726
541 CCGGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 600
3727 GCAAGTGTGAG 3786
601 GCGGTGCTGAG 660
3787 AAGTCCCGGAG 3846
661 AAGTCCCGGAG 720
3847 TCCCGTGTGAG 3906
721 TCCCGTGTGAG 780
3907 CTGACAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAG 3966
781 CTGACAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAG 840
3967 CCGGTGCTGAG 4026
841 CCGGTGCTGAG 900
4027 TACCGAG 4086
901 TACCGAG 960
4087 GTGTTTCTGCTCTCTTCAATCCGACAGCTTGAAGTCCGAGAGAGAGAGAG 4146
961 GTGTTTCTGCTCTCTTCAATCCGACAGCTTGAAGTCCGAGAGAGAGAGAG 1020
4147 CCGGAG 4206
1021 CCGGAG 1080
4207 GATGTGCTGAAG 4266

Db	1081	GACGTGCTGAAGCAGCTACTAGCTGACCTCATTTGATTAAGAACTCGAGAGACAAGAAC	1140
Qy	4267	CTTAAGCTGCTGCTCAGAGAGACCTGACTACGTGCTGAGAAATGCTGACATTTGGTT	4326
Db	1141	CCAAAGTGTGCTCAGAGAGAGAGAGATCGGTGGCCGAGAAATGTTGACATTAATGGTTC	1200
Qy	4327	ACTTCTCTCTCTACAAGTTCTCTCAAGAGTGTGTCTGGAGGCCCTCTTCTCTCTGTTC	4386
Db	1201	ACTTCTCTCTCTACAAGTTCTCTCAAGAGTGTGTCTGGAGGCCCACTTCTCTCTGTTC	1260
Qy	4387	TGTGCCATTCAGACGACGAGATGAGAAAGGGCCCATTAATGCCATCAAGGGCGAGGCCCG	4446
Db	1261	TGTGCCATTCAGACGACGAGATGAGAAAGGGCCCATTAATGCCATCAAGGGAGGCCCG	1320
Qy	4447	TACTCTTGAAGCAGAGCAAAAGCTATCCGCCAGAGATTGACTTCAAAAACCTTGTC	4506
Db	1321	TACTCGCTAAGTAGAGCAAACTCATCTCCAGACAGAGATTGAATTAATAACCTTGTC	1380
Qy	4507	AGCTGTGTCAAGCCAGACATATGCCAAGGCCGAGGTGCCAGTAAGAATCTCAATGT	4566
Db	1381	AGCTGTGTCAAGTCCAGATTAATGTTTAACAGCCCAAGATCCCAAGTAAGATCTCACTGT	1440
Qy	4567	GACACCATCACTCAGGTCAAGAGAGAAATTTCTGATGCCATCTTCAAGAAATGTGCTTGC	4626
Db	1441	GACACCATCACTCAGGTCAAGAGAGAAATCTTAATATCCATCTTCAAGAAATGTGCTTGC	1500
Qy	4627	TCCCAACGGGCCCAAGGTGACGATATGATCTGAGTGGGAGCAAGAAAGTGGGGCAAG	4686
Db	1501	TCCCATAGGCCCAAGGTGACGATATGATCTGAGTGGGAGCAAGAAAGTGGGGCAAG	1560
Qy	4687	ATGATCTTGCAGATGAAAGCATCAACCAACAAATTGAAGTATGTAAGCCGACTGAAC	4746
Db	1561	ATGATCTTGCAGAGACGAAGATATCAACCAACAAATGAGATGTAAGAGAGCTGAAC	1620
Qy	4747	ACACTGGGCCCATCAAGGTGACGATGCTGCTGGTGGGACATTAGTTCCAACAGGTG	4806
Db	1621	ACAGTGAAGCCCATCAAGGTGACGATGCTGCTGGTGGGACATTAGTTCCAACAGGTG	1680
Qy	4807	ACAGCCTTAACAAGAGTGAACCACTCCACCGTCTCCAGAGCCTCAGCAAGTAATATGAA	4866
Db	1681	ACAGCCTTAACAAGAGTGAAGAACACTCCACTGTCTCCAGAGCCTCAGCAAGTAATATGAG	1740
Qy	4867	AACATGATCCGGTCAAGGGGACGGCCGACAGCCCTCGCTCAGGAGCACTTAATATCACT	4926
Db	1741	AACATGATCCGGTCAAGGGGACGGCCCTCAACAGCTCTCGTCCCGAACCCATGATCAAC	1800
Qy	4927	CCTGACCTGAGAGTGAAGTCAAGATGTGGCACTTAATGAAACAACAGAGCAAGAGAC	4986
Db	1801	CCTGACCTGAGAAAGGGGAGTCAAGCTTTGGGATCTGTGGAAGAACATAGAACATGAGAC	1860
Qy	4987	CAGAAAGAGGGGGAGCCGGGGGAGCAAGATGATGTAATCTCACTCAACCGACTCTG	5046
Db	1861	CAGAAAGAGGGGGAGCCGGGGGAGCAAGATGATGTAATCTCACTCAACAGACTATG	1920
Qy	5047	GCCACTAAGGCGACACTCAAGAAATTGTGATGACCTCTTTGAGACCATCTTGACAGC	5106
Db	1921	GCCACTAAGGCGCACCTCAGAAAGTTTGTGATGACCTTAATTCAGAACCATCTTGAGTACG	1980
Qy	5107	GCAACCGATGGCTCGCCCTGCGCCCTCGGCGCATCAAGTACATGTTGACTTCTGGATAG	5166
Db	1981	GCCACCTGTGGCTCCGCTCTGCGCGCTAGCCATCAAGTACATGTTGACTTCTGGATAG	2046
Qy	5167	CAGGCTGTAACATGCGATTCATGACCCGACGTCGCGCATACCTGGAAGACATTCG	5226
Db	2041	CAGGCTGTAAGCACGCGATTCATGACCCGACGTCGCGCACCACTTGAAGAGCAATCGC	2100
Qy	5227	CTGCCCCGAGGTTTGGGTCAACATGATCAAGAACCGCAAGTTGTGTTTGAACATCAT	5286
Db	2101	CTGCCCCCTACGTTTTTGGGTGAATATGATCAAGAACCCGATGTTGTGTTTGAACATCAT	2166
Qy	5287	AAGAACAGCATCAAGACGCTGCTCTCTGTGTGTGCTCAGACCTTCATGGAATCTTGG	5346
Db	2161	AAGAACAGCATCAAGATGCTGCTCTCTCTGTGTGTGCTCAGACCTTCATGGAATCTTGG	2220

	Db	534	TCCACGCTCAGAGAACAACGGGTGGGCAAGAACTGCCCTCCACAAGCTGCTGAATGC	5406
	Qy	5407	GACATCCCAGCTTACCAAGATTGGTGAGAGGTATTACTCAGACATAGAGGAATGCCA	5466
	Db	2221	TCCACTTCGGAGCACCGACTGGGCAAGATTCCCCCTCCACAAGCTGCTTAATGCCAAG	2280
	Qy	5407	GACATCCCAGCTTACCAAGATTGGTGAGAGGTATTACTCAGACATAGAGGAATGCCA	5466
	Db	2281	GACATCCCAGCTTACCAAGAACTGGGTAGAGAGTATTATTACAGACATTGGGAAGATGCCG	2340
	Qy	5467	GCCATCAGAGCAACCAAGACATGAAACGATATCTGCTGAGACAGTCCCGCATGACATGAAT	5526
	Db	2341	GCAATATAGTACCAAGACATGAATGATATCTGGCCGACAGCTGGCGGATGCACATGAAT	2400
	Qy	5527	GAGTTCAACACCAGTAGTGSCACTCTTCAGAGATCTTCTCTTATGTGGGCAAAATACAGCGAG	5586
	Db	2401	GAGTTCAATACGATGATGATGGCTCTCAGAGATCTTCTCTTATGTGGGCAAAATACAGTGGAG	2460
	Qy	5587	GAGATCCTTGGAGCTCTGAGACCAAGATGACCAAGTGTGGGAGACAGAACTGCGCTCAAAA	5646
	Db	2461	GAGATCCTTGGAGCCCTGGACCTGATGACATGACCAAGTGTGGGAGAACTGCGCTTACAAA	2520
	Qy	5647	CTAGAACCAAGTCATAACCCCTCATAGACTTAGACAGCTGAAA	5687
	Db	2521	CTAGAACCAAGTCATACTCTCATAGAGCTTAGACAGCTGAGA	2561
RESULT 3				
AY420221			5564 bp	DNA linear GSS 12-DEC-2003
LOCUS			Mus musculus HCM7151 gene, VIRTUAL TRANSCRIPT, partial sequence,	
DEFINITION			Genomic survey sequence.	
ACCESSION			AY420221	
VERSION			AY420221.1	GI:39776178
KEYWORDS			GSS.	
SOURCE			Mus musculus (house mouse)	
ORGANISM			Mus musculus	
			Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;	
			Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.	
REFERENCE			1 (bases 1 to 5564)	
AUTHORS			Clark,A.G., Glanowski,S., Nielson,R., Thomas,P., Kejariwal,A.,	
			Todd,M.A., Tanenbaum,D.M., Civello,D.R., Lu,F., Murphy,B.,	
			Perriera,S., Wang,G., Zheng,X.H., White,T.J., Smitsky,J.U.,	
			Fertigera,S., Wang,G., Zheng,X.H., White,T.J., Smitsky,J.U.,	
			Adams,M.D. and Cargill,M.	
TITLE			Inferred nonneutral evolution from human-chimp-mouse orthologous	
			gene trios	
JOURNAL			Science 302 (5652), 1960-1963 (2003)	
PUBMED			14671302	
REFERENCE			2 (bases 1 to 5564)	
AUTHORS			Clark,A.G., Glanowski,S., Nielson,R., Thomas,P., Kejariwal,A.,	
			Todd,M.A., Tanenbaum,D.M., Civello,D.R., Lu,F., Murphy,B.,	
			Perriera,S., Wang,G., Zheng,X.H., White,T.J., Smitsky,J.U.,	
			Adams,M.D. and Cargill,M.	
TITLE			Direct Submission	
JOURNAL			Submitted (16-NOV-2003) Celera Genomics, 45 West Gude Drive,	
			Rockville, MD 20850, USA	
COMMENT			This sequence as made by sequencing genomic exons and ordering them	
			based on alignment	
FEATURES			Location/Qualifiers	
source			1..5564	
			/organism="Mus musculus"	
			/mol_type="genomic DNA"	
			/db_xref="taxon:10090"	
gene			<1..>5564	
			/locus_tag="HCM7151"	
ORIGIN				
Query Match			.37.1%; Score 2111.4; DB 29; Length 5564;	
Best Local Similarity			58.5%; Pred. No. 0;	
Matches 3140; Conservative			0; Mismatches 2155; Indels 77; Gaps 9	
Qy		142	GCGAGGGTTTCATTCACCTGGTGTGTGATAGAGACAGAGCATTACTTGGGGGCC	201
Db		151	GACTGGGCTTCATCTCATCTGTGTGTTCACGAACAGACAGAGAGGTATATCTTGGCGCT	210

QY 202 GTCAATCGATTTCACAGCTCTCCAGGACCTGGAAGTCTTGTAACGATGACAGAGG 261
Db 211 GTGAACCGTATTTACAGAGCTGTGCGGGAACTTAACCTGCTTGCGGGCCATGTGACAGGC 270
QY 262 CCGGACGAGAGACACCCAGTGTATACCAACCCGAGTGTGCAGACTGTCAATGAGCC 321
Db 271 CCTGTAAAGACATGGAATGTCTACCAACCCCAAGTAAAGTGTGCCCAATGAT 330
QY 322 CTGACCAACCAACCAATGTCAACAGATGCTCTCTATAGATCTTCAAGAGAAACAGGCTG 381
Db 331 CTGGGACGACAGACAGTCAACAGCTTCTACTGCTGAGCTATGACAGCAACCCGCTG 390
QY 382 ATTGCTGTGGAGACCTGTACCAAGGATCTGCAAGCTGTGAGGCTGAGAGACCTCTTC 441
Db 391 TTGGCTGTGTGCAATGCTCTACAGAGGCACTTGCAAGTCTTGAGGTGAGCAATCTCTTC 450
QY 442 AAGCTGGGAGACCTTATCATTAAGAGAGCACTATCTGTCAAGTGTCAACGAGAGGAGC 501
Db 451 AAGCTGGGAGAGCGGACACCGCAAGAAACACTACCTGTCAAGTGTCAAGAGGAGCTGGC 510
QY 502 TCACTCTTTGAGATGATC-----GTCTCTACAGCAACCTGGATGACAAAGCTTCAAT 555
Db 511 AGCATGGCGGAGTGTCAATGCTGATGCTGGGCAACCGGCAAGGCAAGGCAAGCTCTTCGTG 570
QY 556 GCCACGCGAGTGTATGGGAAGCCGAGATATTTTCCACCATCTTCAGCCGGAACCTGACC 615
Db 571 GGCACATCCATTCATGTCGAATGCAATGAAATATCTCCCACTGTCTCAAGTGGCCGCTCAG 630
QY 616 AAGAACTGTAGGCGGATGCAATGTGCGGATGCTCTCAATATGATGATTCGTGGCTGG 675
Db 631 GCCAACAGAGAAAGTGTGACATGTTGGCTTGTGTACCAAGATGATGTTGTATCATCT 690
QY 676 ATGATTAAGATCCCTTTCGAGACACCTTCAACCATCATCCCTGACTTGTATATCTACTATGTC 735
Db 691 CAGCTTAAGATCCCTTTCAGACATCTGTCCAAATGCTCCGAGCTTGTACATCTACTATGTC 750
QY 736 TATGTTTATAGCAATGCGAACTTGTCTACTTTTGAACCTTCAACCTGAGATGATGTCCT 795
Db 751 TACAGTTTCGAGCGAGCGAGCTGTGTACTACTCTCACTGACGTGAGCAACGACGCTT 810
QY 796 CCACAGAGCTCAACCAAGAGAGAGAGTATATCAATCAAGCTCGTGAAGCTTTGCAAG 855
Db 811 ACCTGCGCGAGCGCGCTGGGAGACCTTCTCACTCAAGATGTGGCGCTGTGCTG 870
QY 856 GAGGACACAGCTTCAACTCTATGTAGAGGTGCCATTGCTGTGAGCGCACTGGAGTGG 915
Db 871 AATGACCCCAAGTCTCTACTCTTACGTGAGATTCCTCCATGCTGTGAGCGCTGGCAGT 930
QY 916 GAGTAACGCTCTGTCGAGGCTGCTTACTGTCCAAACGCGGAGCGCTGTGCAAGAAC 975
Db 931 GAGTATCGCTGTGTGAGAGCGCTTACTGAGCGCGGAGACAGGCTCTGGCAGAGCAG 990
QY 976 CTTGAGTCCATCCAGATGATGACCTGCTCTTCAACGCTCTTTCAGAGGCGCAAGAGCG 1035
Db 991 CTGGGCTGTGTGAGATGAGAGAGTGTGTGTCAACGTTTCGACAGGCGCAGAAAGAC 1050
QY 1036 AAAATTAATCCCTGTGATGATCGGCGCTGTGATCTTTCATCTTTGAAGAGAGTAATGAC 1095
Db 1051 CGCGTGAAGCCACCAAGAGATGCGCGCTGTGCTGTTCACACTGAGGCGCCATCAAGAG 1110
QY 1096 CGCATTAAGAGAGCGGCTGCACTTGTTCACGCGGCGAGGCAACGCTGACCTGGCTGG 1155
Db 1111 AAAATTAAGAGAGCGCATTCATCTGTGTACCGTGTGAAGAGCAAGCTCTCTGCTGG 1170
QY 1156 CTCAAGGTGAAGACATCCCTGTGAGAGTGGCTCTTAACCAATGACATTAATCTGTG 1215
Db 1171 CTGTCTCAACAAGAGCTGGGCTGTGATCACTCGNNNNNNNNNNNNNNNNNNNNNNNNNN 1230
QY 1216 GGCCTGACATGATGCTCCCTGGAGTGTCCGACATGCTGTGAGATTCCTGCTTTC 1275
Db 1231 NNN 1290

QY 1276 ACGGAGGACAGGACCCGATGACGCTGTGATGCGCATATGCTCAAGAAACCACTCTGTG 1335
Db 1291 NNN 1350
QY 1336 GCTTTTGTGGCACCAAAAGTGAGCAAGTGAAGATCCGGGTGATGACCCAGAGGC 1395
Db 1351 NNN 1410
QY 1396 AAC-----GCCCTCAATATGACAGCGTGAAGGTGTGAGACCCCGGCGCAGTCTGC 1446
Db 1411 AGCGGAGCGCGGCTGGCTTATGAAGTGTGTGCTGAGAGGAGAACCTTATCTCTG 1470
QY 1447 CGGATATGCTTCTCCAAAGAACAGAGCACTTATCATATGATGACAGAGCACTC 1506
Db 1471 CGTACCTGTCTCTAGCCCAACCGCAAGTACTTATGCTATGACGAGAGACAGGTG 1530
QY 1507 ACCAGATCTCTGTGAGTCTGTGTGATCATATGAGCTTCCGCGAGTGTCTGCTCA 1566
Db 1531 ACGAGGTGCTGTGAGAAAGCTGTGCAATATCTTCTGTGAGCTGTGTGAGGCTGT 1590
QY 1567 GAGCAACCCCACTGTGCTGTGTGTGTGCTGCAACAGTGTGCAACCCGAAAGAGCGGTGT 1626
Db 1591 CGGACCCCACTGTGCTGT 1650
QY 1627 GAGCGGTCAAGAGACCCCGAGGTTGCTCGAGATGAAGCATGTGTCCGCTGACG 1686
Db 1651 GAGCGGAGAGAGAGCCCAAGCTTGTGCTTCACTGTGAGCTGTGTGTGTGTGTGTGTGT 1710
QY 1687 GTCCATCCCAATATCTCCGTCTCAGTACAGAACGCTGCTGTGTGTGTGTGTGTGTGTGT 1746
Db 1711 GTCCAGCGCGCATGT 1770
QY 1747 AATGTCCGAGCTGTGAGCTGTGCTCAACCTGACCTTTGAGACCTGTGAGATGAT 1806
Db 1771 AATGTGCTGTGACCTGTGAGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1830
QY 1807 GCGCTGT 1866
Db 1831 AGCACTCTGGAAGATGCGCGCATGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1890
QY 1867 ATCATCAAGAGAAATGGGACCAACATGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1926
Db 1891 ATCAAGAGGATGAGGAGACCAAGGAGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1950
QY 1927 GGCATGACCTTGTGCGAGACCAAGCTTGTCTTTCATTAATTCAGAGTGTGTGTGTGTGTGT 1986
Db 1951 GGAAGAAATTTGTGATCTGTGAGCTTGTCTTATTAATTCAGAGTGTGTGTGTGTGTGTGT 2010
QY 1987 CTGTCTGCTGTGAGATGTCAACCGCTGTGCACTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 2046
Db 2011 CTGGAATGTGTCAACGCTGTCTTCCCTGCTGATGTGTGTGTGTGTGTGTGTGTGTGTGTGT 2070
QY 2047 CATGACCCCAAGACTGTCTCTTTCAGAGAAAGCGAGTGAAGCTGTCCGAGAGCTGCCCC 2106
Db 2071 AACATGTGCGGTGATGTGCTTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 2130
QY 2107 CAGCTGTGCGAGTGTGACAAATCTGTGTGCGGTGTGAGTGTGTGTGTGTGTGTGTGTGTGTGT 2166
Db 2131 CAGATCTGTGCTTGTACCAATGT 2190
QY 2167 AAGGCCAAGAACTTCCCGACCGCAAGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 2226
Db 2191 GCTGTACGGAACCTGTGCGACCAACATGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 2250
QY 2227 ATTGAGGAGAGGACAGAGTGTGCGGCTGTGCTTCAACAGCTTCCAGGCTACAGTGC 2286
Db 2251 ATCCAGGACGCGCGCGCGAGT 2310
QY 2287 CAGAACACTCTTATTTCTATGAAGGATGTGAATCAACAACTGTCCGTGTGAGTGTGCA 2346
Db 2311 CAGAACTCTGTGNN 2370
QY 2347 GTGTGTGGAATGGGCACTTCAACATGTGAACCAAGCTCAGATTAAGTTCACCTCTAC 2406

[illegible]

3451	CGGAGCCGCTGTA	CGGAGCCACTTAGG	CGCTTACG	GGCTCTTGGA	GCATGAAGCCAGCTCC	35110
QY	3478	CCCATCATCTTA	AAAGGCAAGAAC	CTGATTC	CCGCGCTGGCTGGGGCAACGTAA	3517
DG	3511	CGCGTATCTCTA	AGGGGCGGAATCTCT	CCACCTGCTCTT	--GGCAATTTCCCGGCTC	3567
QY	3538	AACATCACTGTG	CTGTGGTGGGAGAA	CCGTGAC	CGTGAACGCTGCAGATGTC	3597
DG	3568	AACATCA	CGGTGTCA	TGCGCTCCAC	CACTTGATCTCTCAACCGGTATGAGACACACTG	3627
QY	3598	CTCTGGAGAT	CCCCCAACCTCAT	CCGAGGCA	CAAAATGATGGCCGTGTCGGTGGCATG	3657
DG	3628	CTGTGTAAAGAC	CCCAACCTCA	CGGGGACAGACAA	ATCTACGGTGCGGGCGGGGGCTTC	3687
QY	3658	GAGTACT	CCCCGGGAGATG	GTACATTTG	CCCCGAGACGCCGCTCA	3717
DG	3688	GAGTTCTCA	CTGGGAGATG	CTGACAGTATTT	CGGACAGCTGCTGACACATG	3747
QY	3718	GTACGAT	CGCAATG	GTGGGCGCTC	CTCATATTTATTCGCGGCGTGTCAATTTG	3777
DG	3748	GTGGGTAT	CGGTGGTGGGGGTGG	CTTGTGCTGTG	CTCATCTGCGCTGTGTCTATG	3807
QY	3778	TATTAACGCA	AGTCCCGGAAATG	ATGACCTCA	CGCTGAAGCGGCTGCAGATGACATG	3837
DG	3808	TACAAACGCA	AAATCCGGGATG	CTGACCTGA	ACCTCAACGCGCTGCAGCTCCATG	3867
QY	3838	AACCTGGA	ATGCCGTGTGGCCCTG	GAATGGA	AGGAACCTTTGGCGAGCTGCACAGGAC	3897
DG	3868	AACCTGGA	ATGACGAGGTGGCCCTG	GAATGGA	AGGAACCTTTGGCGAGCTGCACAGGAC	3927
QY	3898	ATCCATGAGCTGA	CCAGTGA	CTGAGATG	AGACCGGATTCCTGTTCTTGATATTA	3957
DG	3928	ATCCATGAGCTGA	CCAGTGA	CTGAGATG	AGTGGTGGATCCCTCTTCTTGATACCGACA	3987
QY	3958	TACACCA	TGGGGGTGCTGTTCC	CAAGAA	TTGAACACACCTGTCTCTCCGGGACCTTGAG	4017
DG	3988	TATGCCAT	GGGAGTGTCTGTTTCT	CTGGGAT	TGGAGCCATCTGTATCTGAGAGATG	4047
QY	4018	GTCCCGGCTA	CCGGACAGAGCG	TGTGAGAA	GAAGCCTTGAACTTTGCTCCAGCTATC	4077
DG	4048	GT-----	GCAGGCCAATGT	GAGAAATG	CGCTGACACCTGTTCCGACACTGTG	4095
QY	4078	AACAA	CAAGGTGTTCC	GTGCTGCTTCA	CCGACCGCTTGATGCTCCAGCTAGCTTCTCC	4137
DG	4096	ACCA	GAAGCATTTCTCTGT	CTTACTTATCTCG	CACATGAGGCCACGATAGCTTCTCC	4155
QY	4138	ATGCGCA	ACCGTGGCA	CGTGGCTCA	CTCATCTACATGACCGTGTGACAGCAAGCTGGAG	4197
DG	4156	ATGCGTAC	CGTGGAAAGTGG	CCCTCACTATCTA	TGACAGCCCTGACAGGTGATGGAG	4215
QY	4198	TAGCCCA	CTGATGTG	CTGAGACAG	CACTGTGGCCACCTCATTTGACAAGACTGGAGAC	4257
DG	4216	TATGCCA	CAGAGATG	CTTAAGCA	TTGTGTGTGACCTCATCGAGAAATTA	4275
QY	4258	AAGAACCA	CCCTTAAGCT	GTGCTCA	GAGAGACTGATGATCATGTGTGAGTGAATGCTGAC	4317
DG	4276	AAGAACCA	CCCTTAAGCT	GTGCTCA	GAGAGACTGATGATCATGTGTGAGTGAATGCTGAC	4335
QY	4318	AATG	GGTTACTTCTCTCTCT	CTACA	AGTTCCTCAAGGATGTGCTGGGAGCCCTCTCTTC	4377
DG	4336	AATCGTT	CACTTCTCTTTTGT	TAATGTTCTCT	CAAGAGTGTGCTGGGAGCCGCTCTTTC	4395
QY	4378	TCCCTGTT	CTGTGCATCA	AGACAGATG	GAAGAGGCCCTCATTTGAGCCCATCA	4437
DG	4396	ATGCTGTA	CTGTGCATCA	AGACAGATG	GAAGAGGCCCTCATTTGAGCCCATCA	4455
QY	4438	GAGGCGGCT	ACTCTTTGAG	GAAGACAA	AGTCACTCCGACAGATTTGACTACAAAC	4497
DG	4456	GAGGCGGCT	ACTCTTTGAG	GAAGACAA	AGTCACTCCGACAGATTTGACTACAAAC	4515
QY	4498	CTG	GTTCTGAGCTGTG	ACCCAGACAT	GTCCCAAGCCCCGAGGTCCAGTAAAGATC	455
DG	4516	CTG	ACCTGGA	CTCGTGAAC	CTTGGAATGCGCCGAGGTTCTGTGAAGGG	457

QY 4558 CTCACATGTCACCATCACTCAGTCAAGAGAGATTCGTGATCCATCTTCAAGAT 4617
DB 4576 CTGAACGTGACACCGTGAACCCAGGTCAAGAGAGAGATCTGTGATCTGTGACAGAGGT 4635
QY 4618 GTGCTTGTCTCCACCGCGCCCAAAAGCTGCAATATGATGTGAGTGGGACAGAGAACT 4677
DB 4636 GTGCTT-----AAGTGGCGCCAGGCTGCC 4660
QY 4678 GGAGCAGAGATGATCTTGCAGAGTGAAGATACCAACCAAGATTGAGATTTGGAAG 4737
DB 4661 ATGGCAGCATCAATCTTCAAGAGAGAGAGATCAACCACTAAATTGACATGATGAGAG 4720
QY 4738 CGACTGAACACCTGGCCCACTACAGAGTCCAGATGTTCCGTGGTGGCATTTAGTTC 4797
DB 4721 AGGCTGAACACACTAGCTCACTATCAGTGAAGAGATGGGTCACTGGTGGCATTAAGTACC 4780
QY 4798 AAGCAGGTGACAGGCTTAAAGCAGTGAACCACTCCAGCTCCAGAGCTCAGAGAT 4857
DB 4781 AAGCAGCATGCGCTTCAAC--ATCTCAATCTCTCTTCACTTAAGTCCCTCAGC 4837
QY 4858 AAATATGAAAATCATGATCCGCTACACGGGACGCGCCGACAGCTCCGCTCAAGACACT 4917
DB 4838 AGATACGAGAGATGCTGCGCTACAGCCAGTAGCCCGACAGCTCAGCTCAGACAGACC 4897
QY 4918 ATGATCACTCTGACCTGAGAGAGTGAAGTGAAGTGGGACCTAGTGAAGAACAGAG 4977
DB 4898 ATGATCAAGCAGCTGAGAGAGTGAAGCAAACTGTGGACCTGTGTGAAGAACAGAG 4957
QY 4978 CACGAGACAG 5037
DB 4958 CACCTGAGACCAAGCTGAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 5017
QY 5038 CGACTCTGCGCATTAAGAGAGACCTGACAGAGATTTGTGATGACCTTTGAGACATC 5097
DB 5018 AGGCTGTGGGACCAAG 5077
QY 5098 TTGACAG 5157
DB 5078 TTGACAG 5137
QY 5158 CTGATGAG 5217
DB 5138 TTGATGAG 5197
QY 5218 AGCATTGCTGCGCCCTGAGGTTTGGGTCAACATGATCAAGAACCGAGTTGTGT 5277
DB 5198 AGCATTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 5257
QY 5278 GACATCAATGAAGACAGATCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 5337
DB 5258 GACATCAAG 5317
QY 5338 GACTTGTCTCAAGTCAAG 5397
DB 5318 NNN 5377
QY 5398 TATGCAAGAGATCCAGCTACAGAGATTGGGTGAAGATTACTACAGATAGAGG 5457
DB 5378 NNN 5437
QY 5458 AAGATCCAGAGCATACGAG 5489
DB 5438 NNN 5469

RESULT 4
AY420666 5520 bp DNA linear GSS 12-DEC-2003
LOCUS Homo sapiens HCM7300 gene, VIRUAL TRANSCRIPT, partial sequence,
DEFINITION genomic survey sequence.
ACCESSION AY420666
VERSION AY420666.1 GI:39776623

KEYWORDS GSS.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
REFERENCE 1 (bases 1 to 5520)
AUTHORS Clark,A.G., Glanowski,S., Nielson,R., Thomas,P., Kejariwal,A.,
Todd,M.A., Tanenbaum,D.M., Civallo,D.R., Lu,F., Murphy,B.,
Ferriterra,S., Wang,G., Zheng,X.H., White,T.J., Smitsky,J.J.,
Adams,M.D. and Cargill,M.
Interfering nonneutral evolution from human-chimp-mouse orthologous
gene trios
Science 302 (5652), 1960-1963 (2003)
JOURNAL 14671302
PUBMED 2 (bases 1 to 5520)
REFERENCE Clark,A.G., Glanowski,S., Nielson,R., Thomas,P., Kejariwal,A.,
Todd,M.A., Tanenbaum,D.M., Civallo,D.R., Lu,F., Murphy,B.,
Ferriterra,S., Wang,G., Zheng,X.H., White,T.J., Smitsky,J.J.,
Adams,M.D. and Cargill,M.
Direct Submission
Submitted (16-NOV-2003) Celera Genomics, 45 West Gude Drive,
Rockville, MD 20850, USA
This sequence as made by sequencing genomic exons and ordering them
based on alignment.
FEATURES
source location/Qualifiers
1..5520
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
<1..5520
/locus_tag="HCM7300"
ORIGIN
Query Match 35.4%; Score 2017.2; DB 29; Length 5520;
Best Local Similarity 55.4%; Pred. No. 0;
Matches 3018; Conservative 0; Mismatches 2399; Indels 39; Gaps 9;
QY 152 TCAATCAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 211
DB 98 TTACCACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 157
QY 212 TTACCACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 271
DB 158 TCTTTAGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGT 217
QY 272 ACAACCCCAAGTGTATACCAACCCCGCATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 331
DB 218 ACAAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 277
QY 332 CCAACATGTCAAG 391
DB 278 TGAACAACATCAAG 337
QY 392 GGAGCTGTACCAAG 451
DB 338 GGAGCATGTGGAG 397
QY 452 AGCTTATATTAAG 511
DB 398 AGCCGACACCAAG 457
QY 512 GAGTATGCTGCTCTCAAG 571
DB 458 GTGTATTTGAG 517
QY 572 GGAAGCCGAGATTTTTCACAG 631
DB 518 GCAAGTCGAGAGATCTTCCACAG 577
QY 632 ATGCAATGTCGAG 691
DB 578 CGACATGTTCACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 637

QY 692 CGAGACCTTCACCATCATCCCTGACTTGTATNTCTACTATGTCTATGTTTATGACAGT 751
 DB 638 CAGACACGCTGCTCTTGTATCCCTGCTTGTGACATCTACTACATCTACGGCTTCTGACGG 697
 QY 752 GCAACTTGTCTACTTTTGTGACCTTCCAACTGAAATGCTGTCTCCACAGCTTCCACCA 811
 DB 698 CCTCTCTCTGTACTTCTCCGACGCTGACGCTGAAACCCAGCAAGCTGTGTCACAG 757
 QY 812 CCAGAGGACGAGGTATATCATCCAGCTGTGAGGCTTGTGCAAGAGGACACAGCTTCA 871
 DB 758 CGGCGCAAGAAATTTTTCAGCTCCAAATGTGTGCGATGTGCGGAGAGCTCAGAGTCT 817
 QY 872 ACTCTATGTAGAGTGGCCCAATGCTGTGAGCGCAGTGGGTGGAGTACCGGCTGTGC 931
 DB 818 ACTCATACGTGGAATTCCTCCATCGGCTGTCTGTGGCGCGGTGAGTACCGGCTGTGC 877
 QY 932 AGGCTGCTTACCTGTCCAAAGCGGGGCGGTGTGTCAGAGCCTTGGAGTCCATCCAG 991
 DB 878 AGAGCGCCACCTTGGCCAAAGCTGTGCTGTCTGTGCGCCAGGCTCTGGCGGTGCGGGCTG 937
 QY 992 ATGATGACCTGTCTTTCACCGTCTTCTCCAAAGGCGCAAGCGGAAATGAAATCCCTGG 1051
 DB 938 ATAGAGACGTCCTTTCACCATCTTCTCTCAGGGCCAGAAAGCCGGGCAAGCCCAAGCC 997
 QY 1052 ATAGAGTGGGCTGTGTCATCTTCAATCTTGAAGCAATAATGACCGCAATTAAGAGCGAG 1111
 DB 998 GGCAGACCATCTCTCTCTCTTTCACCTGACCAATCAATGCTTCAATCCGCGCCGCA 1057
 QY 1112 TGCAGTCTTGTATCCGCGGCGAGGCGACGCTGACCTGACCTGTGCTCAAGTGAAGACA 1171
 DB 1058 TCAGTCTGTCTATCGTGGGGAAGGCACTGTGGCTGTGCTGCTGCTGCTGAACAAGAGGC 1117
 QY 1172 TCCCTGTGACAGTGGCTTAAACATGACATTAATCTGTGGCCCTGAGCATGAATG 1231
 DB 1118 TGCCTGTGACCAACCCCCCATGACATCAACGCAATCTGTGGGCTGTGTGTGAAAC 1177
 QY 1232 CTCCTGTGAGTGTCCGACATGCTGTGTGAATTCCTGTCTTCAAGAGACAGAGGACC 1291
 DB 1178 AGCTCTGTGAGGCTGTGATGTGAGGGGCTGTGCTGTGCGCAGACAGAGGAGC 1237
 QY 1292 GCAATGCTGTGATGCAATATGTCTACAAGAACCACTCTGTGGCTTGTGGGACCA 1351
 DB 1238 GCAATGGCAGCTGTGGCGCTTACACCTACCGCAGACCTGTGTCTTCAATTTGGGACGC 1297
 QY 1352 AAGTGGCAAGCTGAAGAAGATCGGGGTGATGGAACCCAGGGGCAAGCCCTCTCAGTATG 1411
 DB 1298 GCAAGCGGAGCTTGAAGAG---NNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN 1354
 QY 1412 AGAGGCTGACGCTGTGGACCCCGGCACTCTCGGGGATATGGCTTCTCCAGAGACC 1471
 DB 1355 NNN 1414
 QY 1472 ACAGGCAACTCTATCATGTGACAGAGGAGCTCACAGAGTCCCTGTGAGTCTGTG 1531
 DB 1415 NNN 1474
 QY 1532 GTCAATATCAGAGCTGTGGCGAGTGTCTGTGCTCAGGCGACCCCACTGTGTGTGTG 1591
 DB 1475 NNN 1534
 QY 1592 TGCCTGACACACCTGTGACCCCGAAGAGGAGGTGTGAGCGTCCAAAGAGACCCCGAGGT 1651
 DB 1535 TGTGTGACACACAGGTGTGCTGCGCAAGGGGCTGTGTGTGGGCTGTGTGCCCAACAGGCT 1594
 QY 1652 TTTGCTGTGAGATGAAGAGTGTGTGCGGTGACGCTCATGCCAACAATATCTCGTCT 1711
 DB 1595 TTTGCTGTGAGAGCTGAGCAAGTGTGTCCAGGTGTGCGGGTCCGAGCCCAAAATGTGTCAAGTGA 1654
 QY 1712 CTCAGTACACGTGTGTGTGTCTGTGAGAGTCAATGTCTCGGAGACTGTCAAGCTGGG 1771
 DB 1655 GGTCACTGTGGGTGTGAGTGTGACGCTTGTGACACGTCGACAGACTCAAGTGTGGGGG 1714
 QY 1772 TCAACTGTACCTTTGAGGACCTGTCAAGATGATGAG---CTGTGTGTGGGCAATCAGA 1828

DB 1715 TGAAGTGGCTTGTGAGAGGCGCGGAGAAAGAGGCGGTCTGTGCTCCCTCCGGTGAAC 1774
 QY 1829 TCCAGTGTACTTCCCTGTGAGCCAGAGAGGTGCCCGGATCATCAGAGAAATGGGAGCC 1888
 DB 1775 TGTGTGCCCCCTACCTCCCTCCAGAGGTCCAGACTCTTACAGAGGGGCAATGGGGCCA 1834
 QY 1889 ACCATGTCTACAGCTTCAAGCTCAATCAAAAGAGACCGGCAATGACCTTGTGAGACCA 1948
 DB 1835 CCGCACTGTGGGCTGTGAGCTTCTCTCAAGAGAGACGGCTGTGAGGCTTGTGCGGCTG 1894
 QY 1949 GCTTGTCTTCAAAATTTGAGCGTCCACAAATTTGTGCTGTCTGTGCGTGTGAGAGTCAAT 2008
 DB 1895 ACTTTGTCTTCAAACTGTGAGCGTCTCAGTGTGTGCAATGTCTGTGTGGCAGGCTT 1954
 QY 2009 ACCGCTGCACTGTGTAAATACCGGATGTGTGACCCAGTCCCAAGACCTGTGCT 2068
 DB 1955 ACCCTGTGCACTGTGTAAATCTGTGACCGTGTACAGCGGCTCCCAAGAGTCTCT 2014
 QY 2069 TCCAGAGGCGCAGTGAAGTGTGCGAGAGTGTGCGCCAGCTGTGTGAGTGAACAGA 2128
 DB 2015 TCCAGAGGCGCAGTGTGCAAGGCTGTGAGGCTGTGCGCTGAGATCTGTGCGGAGGAGC 2074
 QY 2129 TCTGTGTGCGGTGTGAGGTGTATCAAGCTATCAAGCTGAAGGCGCAAGAACTTCCGAGC 2188
 DB 2075 TCTGTATCCCGTGTGGGTGTATGAGCTTCTTCACTTGTGCGGTGAAGAACTTCACTGAGC 2134
 QY 2189 CCGAGTGTGGGAGCGGTGTGCAAGATGTCTTCAATTCAGGCGAGCGAGCAGAG 2248
 DB 2135 GCGAGTGTGGGCGAGAAATTAATGAGTGTGTGCGGTGTGAGGCGGCGGCGAGGAGG 2194
 QY 2249 TGCCTGCGGTGTGCAAGAGCTGTGAGGCTGTGAGGCTGTGAGGCTGTGAGGCTGTGAGG 2254
 DB 2195 TGCCTGCGGTGTGCAAGAGCTGTGAGGCTGTGAGGCTGTGAGGCTGTGAGGCTGTGAGG 2254
 QY 2309 AAGGATGTGAGTCAAACTGTGAGGCTGTGAGGCTGTGAGGCTGTGAGGCTGTGAGG 2368
 DB 2255 AAGGATGTGAGTCAAACTGTGAGGCTGTGAGGCTGTGAGGCTGTGAGGCTGTGAGG 2314
 QY 2369 ACATTTGACAACTGTGAGTCAAACTGTGAGGCTGTGAGGCTGTGAGGCTGTGAGG 2428
 DB 2315 CCATTTGACAACTGTGAGTCAAACTGTGAGGCTGTGAGGCTGTGAGGCTGTGAGG 2374
 QY 2429 GCTGTGCGGTGTGCAAGAGCTGTGAGGCTGTGAGGCTGTGAGGCTGTGAGGCTGTGAGG 2488
 DB 2375 GCTGTGCGGTGTGCAAGAGCTGTGAGGCTGTGAGGCTGTGAGGCTGTGAGGCTGTGAGG 2434
 QY 2489 GCGAGTGTGCAAGGCTGTGAGGCTGTGAGGCTGTGAGGCTGTGAGGCTGTGAGGCTGTGAGG 2548
 DB 2435 ACAGTGTGAGTGTGAGGCTGTGAGGCTGTGAGGCTGTGAGGCTGTGAGGCTGTGAGG 2494
 QY 2549 GTGCCAAAGCAAGTGTGCAAACTGTGAGGCTGTGAGGCTGTGAGGCTGTGAGGCTGTGAGG 2608
 DB 2495 AGAAGGCGACCGGTGTGAGGCTGTGAGGCTGTGAGGCTGTGAGGCTGTGAGGCTGTGAGG 2554
 QY 2609 GGAAGGCGGCGACCAAGTGTGCAAACTGTGAGGCTGTGAGGCTGTGAGGCTGTGAGGCTGTGAGG 2668
 DB 2555 AGAAGGCGGCGACCGGTGTGAGGCTGTGAGGCTGTGAGGCTGTGAGGCTGTGAGGCTGTGAGG 2614
 QY 2669 TGCCTTCCATGTCAAGTGTGAGGCTGTGAGGCTGTGAGGCTGTGAGGCTGTGAGGCTGTGAGG 2728
 DB 2615 TGGGCTGTGAGGCTGTGAGGCTGTGAGGCTGTGAGGCTGTGAGGCTGTGAGGCTGTGAGG 2668
 QY 2729 CTGCAAGACAGTGTGAGGCTGTGAGGCTGTGAGGCTGTGAGGCTGTGAGGCTGTGAGGCTGTGAGG 2785
 DB 2669 GTGTGTGAGGCTGTGAGGCTGTGAGGCTGTGAGGCTGTGAGGCTGTGAGGCTGTGAGGCTGTGAGG 2728
 QY 2786 TGTGTGAGGCTGTGAGGCTGTGAGGCTGTGAGGCTGTGAGGCTGTGAGGCTGTGAGGCTGTGAGG 2845
 DB 2729 CCGTGTGAGGCTGTGAGGCTGTGAGGCTGTGAGGCTGTGAGGCTGTGAGGCTGTGAGGCTGTGAGG 2788
 QY 2846 ATTACTTATGACTGACTCTCTCAATGTGAAGCCAGGCGGGGCGCAATGTCCGAG 2905

Db 2789 ACAGCTTTGTGACCCCAAGCTTTGACCAAGTGAAGTCCAGCCGCTGCGGCGTCCGGG 2848
Qy 2906 GGACCCCAAGTGAACATCAAGGCAACCACTGAATGCGGAGCAAGTGTGTATGT 2965
Db 2849 GCACACGCTTACCATCTCAGGCGACTCTGTGATGTGGACAGGGCTCACAGTACTG 2908
Qy 2966 TTGGAAGCAGACCTGTCTCTTCCACAGGCGATCTCCATCTTCAATGTCTGCAACCA 3025
Db 2909 TGAGGACAGGAGATGCGAGTTGTGTAGAGAGATGCAAGGGAGATGTGTGATCTCAC 2968
Qy 3026 CATCTCAATGATGAGTGTAG--AGATGAAGTGTGTGTGAGTGTGACAGGGCCACA 3082
Db 2969 CTCTCTCCACCTTGCGGCCCGACGCGCCCATCACTGTCATTTGACCGGGCTAACA 3028
Qy 3083 TCCA---CCAGACCTGGTCTTTCAATGTGAAGACCCACCATGTGCGGATTGAG 3139
Db 3029 TCTCCAGCCCCGGGCTCATCTACACTCACTAGGACCCCATCCGCTTGAAG 3088
Qy 3140 CAGATGAGCATTTGTCAGTGGAACAACCCATGCGCTATGGGGACCCACCTGGACC 3199
Db 3089 CCACCTGGAGCATCATCAANN 3148
Qy 3200 TCATACAGAACCCCAAGTCCGTGCCAAGCATGGAGGAAGAGCATCATATATCTGTG 3259
Db 3149 NNN 3208
Qy 3260 AGGTTCTGAACGCTACTGATGAATGACTGTCAAGCGCCGCCCTCGCTCTGGGCTTGAAC 3319
Db 3209 NNN 3268
Qy 3320 ACCAGTCAGACTCAACGAGAGGCCCAAGAGTTTGGCTTCATCTGGAACAAGTCCAGT 3379
Db 3369 NNN 3328
Qy 3380 CCCTGCTCATCTCAACAAGCAACTTCACTATCCCAACCCGGTGTGGAGCCT 3439
Db 3329 NNN 3388
Qy 3440 TTGGTCCCTCAGGAATCCTGAGACTCAAGCTGCGACGCCCATCATCTTAAAGGCAAGA 3499
Db 3389 NNN 3448
Qy 3500 ACCTGATCCGCGCTGTGGCTGAGGCAACGTGAAGTGAACCTACTGCTGTGGAGG 3559
Db 3449 NNN 3508
Qy 3560 AGAAGCCGTGACCGGTGACCGTGTCAATGTCAAGCTGCTGTGGAAGTCCCAACTCA 3619
Db 3509 NNN 3568
Qy 3620 TCGGCAAGCAACAAGTGAATGCGCGGTGCGTGGGATGAGTACTCCCGGGAGTGTGT 3679
Db 3569 NNN 3628
Qy 3680 ACATGCCCCGAGACGCCGCTCAAGCCTGCGCCGATGCTCAAGCATGCAAGTGTGGCG 3739
Db 3626 ACATCTGGGAGAGGGGCGCTGACCTACCGGCAATGATGGGCTGCGGGCGGGGCTG 3685
Qy 3740 GCTCTCTCATATTTTCAATGCGCGGTGCTCAATTTGCTATAAAGCAAGTCCCGGAAA 3799
Db 3686 GAGTCTGTGCTGTGGCATCAAGCCGCTGTGTGGGCTACAAACGGAACCTAAGAGCG 3745
Qy 3800 GTGACCTCAAGCTGAAGCGCTGAGATGAGATGCAACCTGAGTCCCGTGTGGCC 3859
Db 3746 CGAGCCGTAACCTCAAGGTCTGAGCTGAGATGAGCAACCTGAGTCCCGTGTGGCC 3805
Qy 3860 TGGAGTGAAGAAAGCTTTTCCAGAGTGTGAGAGAGATCAATGAGCTGACAGTGAAC 3919
Db 3806 TGGAGTGAAGAAAGCTTTTCAAGAGTGTGAGAGAGATCAATGAGCTGACAGTGAAC 3865
Qy 3920 TGGATGAGAGCGGAGTTCGTTCTGTGACTATGAACCTTACACATCGGGTGTGTTCC 3979
Db 3866 TGGAGAGGTGCAATGCCCTTCTGTGACTACCGGACTTACGCCGTGCGGTGTTCTTCC 3925

Qy 3980 CAGGAATTGAAGAACCAACCTGTCTCCGGAGACTTGAAGTCCCGGCTAACCGGAGAGC 4039
Db 3926 CGGGCATGAGAGCCCAACCGGTGTCTCAAGAGCTGG-----ATTNNNNNNNNNN 3973
Qy 4040 GTGTGAAGAAAGGCTGAAGCTCTTCCGCCAGCTCATCAACAAGTGTCTCTGCTGT 4099
Db 3974 NNN 4033
Qy 4100 CCTCATCGACGCTTGAATGCCAGCTAGCTTCTCATGCGGACCGGTGCAACGTGG 4159
Db 4034 NNN 4093
Qy 4160 CCTCATCATGATGACCGTGTGAGAGCAAGCTGAGTACGCCATGATGTGCTGAAGC 4219
Db 4094 NNN 4153
Qy 4220 AGCTGTGCGGACCTCATTTGACAAGACCTGAGAGCAAGAACCCCTAAGCTGTGC 4279
Db 4154 NNN 4213
Qy 4280 TCAGAGAGACTGAGTCAAGTGTGAGAAATGCTGCAATTTGTTTACTTCTCTCT 4339
Db 4214 NNNNNNNNGACAGAGTCAATGTGTGAGAAATGCTTACCACTGGTTCAGTTCCTGCTGC 4273
Qy 4340 ACAAGTCTCTCAAGAGTGTGTGGGGAGCCCTCTTCTCCCTGTCTGTGCCATCAAGC 4399
Db 4274 ATAAATTTCTGAAGAGTGTGTGGGAGCCCTCTCTCTCTTCACTGTGCATCAAGC 4333
Qy 4400 AGCAGATGAAGAGGCCCCCAATGACGCAATCAAGGAGAGCGCGCTACTCTTGAAGC 4459
Db 4334 AGCAGATGAAGAGGCCCCCAATGATGATGCAACAGGGAGAGAGCATCTCCGAGCG 4393
Qy 4460 AGGACAGCTCATCTCGCAGAGATGATGACTCAAAACCTGTGTCTGTAGCTGTCAAGC 4519
Db 4394 AGGACAGCTCATCTCGTACAGAGATGATGACTCAAAACCTGTGTCTGTAGCTGTGC 4453
Qy 4520 CAGACATGCCAACAGCCCCGAGGCTCCAGTAAGATGCTCAACTGTGACACCATCACTC 4579
Db 4454 CGGAGAAAGAGGCGAGCCCGCAGGCTCCAGTGAAGTGTCTCACTGTGACAGCATCACCC 4513
Qy 4580 AGTCAAGAGAAAGTCTGATGCTCAATGCTTCAAGATGTGCTTGTCCACCGGCCCA 4639
Db 4514 AGGCCAAAGATTAAGTGTGTGACACTGTGTCAAGGGCAATTCGTAACCGAGGTCCCA 4573
Qy 4640 AAGCTGCAATTAAGTGTGATGAGTGTGAGAGGAGAGAGGAGCAAGATGATCTTGCAGG 4699
Db 4574 AAGCTGAGAGATGAGCTGAGAGTGTGAGGAGGAGGAGGAGGAGTATGCTGATCTCCAGG 4633
Qy 4700 ATGAAGATCATCAACCAAGATTGAATGATTTGGAACGACTGAACACTGTGCCACT 4759
Db 4634 ATGAGAGATGTCAACCAAGATGAGTGTGATGAGGAGGAGGAGGAGGAGGAGGAGG 4693
Qy 4760 ACCAGTGTCAATGTTTCCGTTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAG 4819
Db 4694 ACCAGTGTCAACAGTGTCTGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAG 4753
Qy 4820 CAGTGAACATCTCAAGCTGTCTCAAGAGCTCAAGAGTGAATATGAAAAATGATCCGT 4879
Db 4754 TGGCCAACTCTTCACTTCACTTCACTCCG---CTCCCTCAGCGGCTACAGAGCTTGTCCGCA 4810
Qy 4880 ACAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 4939
Db 4811 CGGCGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 4870
Qy 4940 GTGAGTCAAGATGTGAGCACTTGAAGAAACCAAGAGAGGAGGAGGAGGAGGAGGAGG 4999
Db 4871 CAGGACCAAAATTTGTGAGCACTGTGAGAAACCAAGAGCAATGCGGAGGAGGAGGAG 4930
Qy 5000 ACCGGGGAGCAAGATGTTGTGTAATCTTACCTGACCCGAGCTCTGTGCTAAGAGCA 5059
Db 4931 ACCGTGAGCAAGATGTTGTGTAATCTTACCTGAGATCTGAGCAAGGCTGTGAGCAAGAGGCA 4990

QY 5060 CACTGACAAAGTTTGTGATGACCTCTTTGAGACCATCTTCAGACGAGCAACCGTGGCT 5119
Db 4991 CACTGACAAAGTTTGTGATGACCTCTTTGAGACCATCTTCAGACGAGCAACCGTGGCT 5050
QY 5120 CTGCGCTGCGCTGCGCTGCGCTGCGCTGCGCTGCGCTGCGCTGCGCTGCGCTGCGCT 5179
Db 5051 CCGGCTGCGCTGCGCTGCGCTGCGCTGCGCTGCGCTGCGCTGCGCTGCGCTGCGCT 5110
QY 5180 ATGGCAATCATGACCGGACGCTGCGCTGCGCTGCGCTGCGCTGCGCTGCGCTGCGCT 5239
Db 5111 GCGAGATCAGGACCGGACGCTGCGCTGCGCTGCGCTGCGCTGCGCTGCGCTGCGCT 5170
QY 5240 TTTGGGTCAACATGATCAAGAACCGGACGCTGCGCTGCGCTGCGCTGCGCTGCGCT 5299
Db 5171 TCTGGGTGAATGATCAAGAACCGGACGCTGCGCTGCGCTGCGCTGCGCTGCGCT 5230
QY 5300 CAGACGCTGCGCTGCGCTGCGCTGCGCTGCGCTGCGCTGCGCTGCGCTGCGCTGCGCT 5359
Db 5231 CCGATGCTGCGCTGCGCTGCGCTGCGCTGCGCTGCGCTGCGCTGCGCTGCGCTGCGCT 5290
QY 5360 ACCGCTGCGGCAAGACCTGCGCTGCGCTGCGCTGCGCTGCGCTGCGCTGCGCTGCGCT 5419
Db 5291 ACCGCTGCGGCAAGACCTGCGCTGCGCTGCGCTGCGCTGCGCTGCGCTGCGCTGCGCT 5350
QY 5420 ACAAGATTTGGGTGAGAGAGTATTACTGAGACATGAGGAAGATGCGACCGCTGAGGAC 5479
Db 5351 ACAAGAGCTGGGTGAGAGAGTATTACTGAGACATGAGGAAGATGCGACCGCTGAGGAC 5410
QY 5480 AAGACATGAGACCATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 5539
Db 5411 AAGACATGAGACCATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 5470
QY 5540 TGAATGCACTTCAGAGATCTTCTCTATGTGGGCAATATACGACGAGAG 5589
Db 5471 TGAATGCACTTCAGAGATCTTCTCTATGTGGGCAATATACGACGAGAG 5520

RESULT 5
AY420668 5447 bp DNA linear GSS 12-DEC-2003
LOCUS AY420668 Mus musculus HCM7300 gene, VIRAL TRANSCRIPT, partial sequence.
DEFINITION Mus musculus HCM7300 gene, VIRAL TRANSCRIPT, partial sequence.
ACCESSION AY420668
VERSION AY420668.1 GI:39776625
KEYWORDS GSS.
SOURCE Mus musculus (house mouse)
ORGANISM Mus musculus

REFERENCE
AUTHORS Clark, A.G., Glanowski, S., Nielson, R., Thomas, P., Kejarival, A.,
1 (bases 1 to 5447)
Todd, M.A., Tanenbaum, D.M., Civeleo, D.R., Lu, F., Murphy, B.,
Ferrera, S., Wang, G., Zheng, X.H., White, T.J., Smitsky, J.J.,
Adam, M.D. and Cargill, M.
Inferring nonneutral evolution from human-chimp-mouse orthologous
gene trios
Science 302 (5652), 1960-1963 (2003)

JOURNAL
PUBMED 14671302
REFERENCE 2 (bases 1 to 5447)
AUTHORS Clark, A.G., Glanowski, S., Nielson, R., Thomas, P., Kejarival, A.,
Todd, M.A., Tanenbaum, D.M., Civeleo, D.R., Lu, F., Murphy, B.,
Ferrera, S., Wang, G., Zheng, X.H., White, T.J., Smitsky, J.J.,
Adam, M.D. and Cargill, M.
Direct Submission
Submitted (16-NOV-2003) Celera Genomics, 45 West Gude Drive,
Rockville, MD 20850, USA
This sequence as made by sequencing genomic exons and ordering them
based on alignment.

TITLE
JOURNAL
COMMENT
FEATURES
Source
1. 5447
/organism="Mus musculus"
/mol_type="genomic DNA"
/db_xref="taxon:10090"

gene
ORIGIN
Query Match 31.1%; Score 1767.2; DB 29; Length 5447;
Best Local Similarity 53.4%; Pred. No. 0;
Matches 2910; Conservative 0; Mismatches 2427; Indels 115; Gaps 9;
<1..>5447
/locus_tag="HCM7300"

QY 150 TTTCATCACTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 209
Db 99 TCGATCACTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 158
QY 210 GATTTCAGAGCTCTCCAGGACCTGAGAGCTTTGGTGAAGCATGAGACAGGAGCCGAGACA 269
Db 159 AGTTTCAGAGCTCTCCAGGACCTGAGAGCTTTGGTGAAGCATGAGACAGGAGCCGAGACA 218
QY 270 GAGCAACCCAGAGTGTATTAACCCAGCCGAGAGCTGTCAGACCTGCAATGAGCCCTGACAC 329
Db 219 GAGCAATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 278
QY 330 CACCAACATGTCAACAAAGTGTCTCTCAATGATCAAGAGAAACAGGCTGATTTGCCG 389
Db 279 TGTGCAATGTGAAACAAAGTGTCTCTCAATGATCAAGAGAAACAGGCTGATTTGCCG 338
QY 390 TGGAGAGCTGTACCAAGGACCTGCAAGCTGCTGAGAGCTGAGAGACCTTCAAGCTGGG 449
Db 339 TGGAGAGCTGTACCAAGGACCTGCAAGCTGCTGAGAGCTGAGAGACCTTCAAGCTGGG 398
QY 450 GAGAGCTTATCATAGAGAGAGACATATGTGTGATGATGATGATGATGATGATGATGATG 509
Db 399 TGAAGCCCAACATGCAAGAGAGACATATGTGTGATGATGATGATGATGATGATGATGATG 458
QY 510 TGAAGTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 569
Db 459 TGAAGTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 518
QY 570 TGGAGAGCCGAGTATTTTCCAGCATCTCCAGCCGAGAACTGACCAAGAACTTGAGGC 629
Db 519 TGGAGAGCCGAGTATTTTCCAGCATCTCCAGCCGAGAACTGACCAAGAACTTGAGGC 578
QY 630 GATGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 689
Db 579 TGAAGTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 638
QY 690 TGGAGACCTTCACCATCATCTCCAGCTTGTATCTATCTATCTATCTATCTATCTATCTAT 749
Db 639 TGGAGACCTTCACCATCATCTCCAGCTTGTATCTATCTATCTATCTATCTATCTATCTAT 698
QY 750 TGGCACTTTGCTACTTTTGAACCTTCAACCTGAGATGAGTGTCTCCACAGGCTCCAC 809
Db 750 TGGCACTTTGCTACTTTTGAACCTTCAACCTGAGATGAGTGTCTCCACAGGCTCCAC 758
QY 699 TGGCTCTTGTGTACTCTTCAAGTGTGAGCTGAGCTGAGACCCAGACAGAGCTGTGGATAC 758
Db 759 TGGCTCTTGTGTACTCTTCAAGTGTGAGCTGAGCTGAGACCCAGACAGAGCTGTGGATAC 718
QY 870 CAACTCTATGTAGAGTGGCCCATTTGGCTGTAGAGGAGTGGGGTGAATACCGCTGCT 929
Db 870 CAACTCTATGTAGAGTGGCCCATTTGGCTGTAGAGGAGTGGGGTGAATACCGCTGCT 878
QY 819 CTACTCTTATGTAGAGTGGCCCATTTGGCTGTAGAGGAGTGGGGTGAATACCGCTGCT 989
Db 819 CTACTCTTATGTAGAGTGGCCCATTTGGCTGTAGAGGAGTGGGGTGAATACCGCTGCT 878
QY 930 GAGAGCTGCTTACCTGCAAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 989
Db 879 GAGAGCTGCTTACCTGCAAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 938
QY 990 AGATGATGACCTGCTTCAACCGCTTTTCAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1049
Db 939 GAGAG 998
QY 1050 GATGAGTGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1109
Db 999 TGGGACACATCTTGTGCTTTTCAACCGCTTCAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1058
QY 1110 GCTGAGCTTGTGTTACCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1169

[illegible][illegible]


```
QY 3318 CCACAGTCAGACCTGACCGAGAGGCCGAGAGATTGGCTTCATCTGGAACAACGTCCA 3377
Db 3194 NNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN 3253
QY 3378 GTCCCTGCTCATCTCAACAAGCACTTCACTACTATCCCAACCCGGGTGTTGAGGC 3437
Db 3254 NNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN 3313
QY 3438 CTTTGGTCCCTCAGAAATCCTGAGCTCAAGCTGGACGCCCATCATCTTAAAGGCA 3497
Db 3314 NNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN 3373
QY 3498 GAACTGATCCGCCCTGTGGCTGGGGGCAAGTGAAGTGAACCTACCTGTGTGTGG 3557
Db 3374 NNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN 3433
QY 3558 GGAGAACCGGTGACCGGTGACCGTGTCAAGTGTCAAGCTGCTCTGAGATCCCCAACT 3617
Db 3434 NNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN 3493
QY 3618 CATCGCAGGCAACAAGTATGCCCCTGTCTGGTGGCATGAGATCTCCCGGGGATGGT 3677
Db 3494 NNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN 3556
QY 3678 GTACATTGCCCGGACAGCCCGCTCAGCTGCCCGCATCTCAGCATCGAGTGGCTGG 3737
Db 3551 GCATCATCTGCTGATCGGGCACTGACCTTACAGCTAGTGGGGCTAGAGCAGGGGG 3610
QY 3738 CGGCTTCCTCATTTTCACTGTGGCCGTCTCATGTGCTTAAACGCAAGTCCCGGA 3797
Db 3611 CGGGCTCTTGTGCTGGCATCACTGTGTGTGGCTTCAAGCGTAAACCTCAAGGA 3670
QY 3798 AAGTGAACCTCAAGCTGAGGGCTGAGATGAGATGAGCAACTGAGTCCCGTGGGC 3857
Db 3671 TGGAGCCGCACTTAAGCGCTTACGCTTACAAATGAGCAACTGGAGTCTCGTGTGGC 3730
QY 3858 CTTGAGTGAAGAAACCTTTGCGAGCTGACAGCGACATCCATGAGCTGACCAAGTGA 3917
Db 3731 CTTGAGTGAAGAAACCTTTGCTGAGCTGAGCTGATATCAATGAATCAACAACA 3790
QY 3918 CCGAGTGAAGCGGATTCCTGCTGAGCTATGAATCAATCAATGAGCTGCTGT 3977
Db 3791 CATGATGTGTCAATCCCTCTTCTTGAATCGGACCTATGCTGTGGGTCTTT 3850
QY 3978 CCCAGAAATTGAAGCAACCTGTCTCCGGGACCTTGAAGTCCGGGCTACCGGAGGA 4037
Db 3851 CCTGGCATGAGGCTCAACCAGTGTCTCAAGGAGCTG-----ATNNNNNNNN 3898
QY 4038 GCGTGTGAGAAAGGCTGAAGCTTCGCGCCAGCTCATCAACAAGAGTGTCTGT 4097
Db 3899 NNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN 3958
QY 4098 GTCTTCATCCGACGCTTGAATCCAGCGTACCTTCTCATGCGGACCGTGGACAAGCT 4157
Db 3959 NNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN 4018
QY 4158 GGCTCACTCATATGACCGTGTGCGAGCAAGCTGAGTACGCCACATGTTGTGA 4217
Db 4019 NNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN 4078
QY 4218 GCAGCTGTGCGCAGACTCATTTGAACAAGACTGAGAGCAAGAACCACTTAAGCTGT 4277
Db 4079 NNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN 4138
QY 4278 GCTCAGAGAGACTGAGTCAAGTGGCTGAGAAAGATGCTGAACCAATTGGTTACTTCTCT 4337
Db 4139 NNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN 4198
QY 4338 CTACAAGTCTCTAAGAGAGTGTCTGGGAGGCCCTTCTCTCTGTTCTGTGGCATCAA 4397
Db 4199 GCATTAATCTCTGAAGAGTGGCGCGGGAGGCCACTTCTCTGTGACTGAGGCATCAA 4258

QY 4398 GCAGCAGATGAGAAAGGCCCATTTGAGCCCATCAAGGGGAGAGGCCCGCTACTCTTGG 4457
Db 4259 GCAGCAGATGAGAAAGGCTCTATTGATGCATTAACAGGCGAGGCCCTTACTCTTAAAG 4318
QY 4458 CGAGCAAGCTCATCCGCGAGAGATTGACTAACAACCTGTGTCTGAGCTGTGTCAG 4517
Db 4319 TGAGCAAGGCTCATCCGCGAGAGATCGACTATAAAGCCCTGACTCTGATTTGTGTGG 4378
QY 4518 CCCAGCAATGCCAAGAGCCCGAGGTCCTCAAGAAATCCCACTGAGACATGAC 4577
Db 4379 CCCGAGAGGAGGAGATGCCCGAGTCCCTGTGAAGGTTCTTAACTGTGAATGATCAC 4438
QY 4578 TCAGTCAAGAGAAATTCCTGATCCACTTCAAGAAATGTGCTGTGCCACCGGCG 4637
Db 4439 CAGGCGCAAGACAAAGTGTGATCTGTGTCAAGGGTATTCATCTCAAGCGCC 4498
QY 4638 CAAAGCTGAGATATGATCTGAGTGGCGACAAGAAATGGGGGCAAGATATCTTGA 4697
Db 4499 CAAAGCTGAGAGATGATTTGAAATGGCGCGAGGCCCATGGCCGATCATCTCCA 4558
QY 4698 GATGAAGACATCAACCAAGATTGAGATGATGAAAGCACTGAACAACCTGGCCA 4757
Db 4559 AGATGAGATCATCAACAAGATTGATGTGATGAGAGAGGATCAATCTGAGCCCA 4618
QY 4758 CTACCAAGTCCAGATGTTCCGTGTGGCATTAATGTCTCAAGCAAGTGAACCTTAA 4817
Db 4619 CTACCAAGTACTGATGTGTTCTTAAATGACATGGTCCCAACAAGTGTCTGCTATTA 4678
QY 4818 CGAAGTGAACAATCCACGCTCCAGAACCTCAGCAAGTAAATATGAACAATGATCCG 4877
Db 4679 CATGGCAATCTGTGATCTTACCTTACCCG---TCACTTGTGCTAGAGAGCTGTGCTCG 4735
QY 4878 GTACACGGGAGGCCCGGACAGCTCCGTGCTCAAGCAACTATGATCACTCTGACCTGA 4937
Db 4736 TGCTGCCAGAGCCCGGATGAGCTCCGTTCCAGAGCACTATGCTACCGCTTACCAAGA 4795
QY 4938 GAGTGAAGTCAAGATGGGACCTAGTGAAGAAACAAGCAAGAGCAAGAGAGAGG 4997
Db 4796 GGCAGTTACCAAGTGTGGCACTGTGAGGAACCAAGCAACATGATCAACGAAAG 4855
QY 4998 GGAACGGGGAGGCAAGTGTGTGAATCTAAGTCCGACCTCTGGCCACTAAGG 5057
Db 4856 AGACCGCGGAGCAAGTGTGTCAAGAAATATCTCAAGGCTCTGCGCAACAAGG 4915
QY 5058 CACATGCAAGATTTGTGATGACCTCTTGAACATCTTGAACGCGCACACGCTGG 5117
Db 4916 CAGCTTCAGAAATTTGTGATGATGATCTTTTGAATCTGTGTATGATGACCAACCGGG 4975
QY 5118 CTGTGCCCTGCCCCGTGCGCATCAAGTACATGTTGAATCTCTGATGAGCAGCTGATA 5177
Db 4976 CTGAGCTTACCTTGGCCATCAAGTACATGTTGAATCTCTGATGAAAGAGCTGACCA 5035
QY 5178 ACATGATTCATGACCCGACGCTCGGCATTAAGTGAAGAAATGAGTGGCTGCGCTGAG 5237
Db 5036 GCGCAGATGATGACCTGATGATGCTGACCACTGGAAGCAACTGCTTACTCTGCG 5095
QY 5238 GTTTGGTCAACATGATCAAGAAACCCGAGTTGTGTTGATCATCATCAAGCAATCCAG 5297
Db 5096 TTTCTGGTGAATGTATCAAGAAATCGCAATTTGTGTTGATCATCAAGAAATAGCAT 5155
QY 5298 CACAGAGCTGCTCTCTGTGTGTGCTGACACTTCAATGATGATCTGCTCCAGCTGAG 5357
Db 5156 CACAGAGCTGTTGTGAGTGTGTGCGACCTTATGATGATCTGCTGTTCTCATATGAG 5215
QY 5358 GACCGGCTGGGCAAGACTCGCCCTCAACAAGCTGCTATGCAAGCAATCCAG 5417
Db 5216 GACCGGCTGGGCAAGAGTCAACCTTCAACAAGCTCCTCATGCAAGAAATATCCAA 5275
QY 5418 CTACAAGATTTGGTGAAGATTTTACTAGACATAGGAAGATGCAAGCAATCAAGCA 5477
Db 5276 TTACAAGCTGGTGAAGATTTACTATGAGATATGCAAGATGAGCAATCAATCAAGTGA 5335
QY 5478 CCAAGACATGAACCATACTGCTGAGCAGTCCCGGATGACATGAATGATTCACAC 5537
```

Db 5336 CAGAGACATGACGCTACTTAGTGAGACAGACCTCCGCTCATGCTAATGCTTCAATGT 5395
Qy 5538 CATAGTGCATCTCAGAGATCTTCTCTATGTGGGCAATATACGAGAG 5589
Db 5396 CTAATGTCACCTCAGCGAGCTTCTTCTATGTACCAAGTACCGTCAAGAG 5447

RESULT 6
AY420667 5520 bp DNA linear GSS 12-DEC-2003
LOCUS AY420667
DEFINITION Pan troglodytes HCM7300 gene, VIRUAL TRANSCRIPT, partial sequence,
genomic survey sequence.
ACCESSION AY420667
VERSION AY420667.1 GI:39776624
KEYWORDS GSS.
SOURCE Pan troglodytes (chimpanzee)
ORGANISM Pan troglodytes
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Pan.
REFERENCE 1 (bases 1 to 5520)
Clark,A.G., Gnanowsky,S., Nielson,R., Thomas,P., Kejarival,A.,
Todd,M.A., Tanenbaum,D.M., Civallo,D.R., Lu,F., Murphy,B.,
Ferreira,S., Wang,G., Zheng,X.H., White,T.J., Shinsky,J.J.,
Adams,M.D. and Cargill,M.
Inferring nonneutral evolution from human-chimp-mouse orthologous
gene trios
Journal Science 302 (5652), 1960-1963 (2003)
PUBMED 14671302
JOURNAL 2 (bases 1 to 5520)
Clark,A.G., Gnanowsky,S., Nielson,R., Thomas,P., Kejarival,A.,
Todd,M.A., Tanenbaum,D.M., Civallo,D.R., Lu,F., Murphy,B.,
Ferreira,S., Wang,G., Zheng,X.H., White,T.J., Shinsky,J.J.,
Adams,M.D. and Cargill,M.
Direct Submission
Submitted (16-NOV-2003) Celera Genomics, 45 West Gude Drive,
Rockville, MD 20850, USA
COMMENT This sequence as made by sequencing genomic exons and ordering them
based on alignment.
FEATURES
Location/Qualifiers
1..5520
/organism="Pan troglodytes"
/mol_type="genomic DNA"
/db_xref="taxon:9598"
<1..>5520
/locus_tag="HCM7300"

ORIGIN
gene
Query Match 29.5%; Score 1681.4; DB 29; Length 5520;
Best Local Similarity 46.8%; Pred. No. 0;
Matches 2552; Conservative 0; Mismatches 2859; Indels 39; Gaps 9;

Qy 152 TCATACCTGCTGCTGATGAGAGACAGACATTTACTTGGGGCCCTCAATCGA 211
Db 98 TTACCACTGCTGCTGACACCGGATGACTGGGAGTGTNNNGGNNNNNNNNNN 157
Qy 212 TTACCACTGCTGCTGACACCGGATGACTGGGAGTGTNNNGGNNNNNNNNNN 271
Db 158 TCTTTAAGNTGGCCCCCACTGACTGAGCTGGGGCCCATGTCAGGGGCCGTTNNNG 217
Qy 272 ACAACCCAGAGTTTACCAACCGGATGCTCCAGACTGAGTGGCCCTGACACCA 331
Db 218 ACAACGTCGCTCTACCGCCGCCAGATGCGCTGTGTGCGCCACCGCTGGCCCG 277
Qy 332 CCAACATGTCACAGATGCTCTCATAGACTCAAGAGAAAGAGCTGATGCTGTG 391
Db 278 TGACAAACATCAACAGCTGCTCTCATAGACTATGCGGCCCGCTGGGCTGG 337
Qy 392 GAGACCTGTACCAAGATGCTGCAAGCTGAGGCTGAGAGACCTTCAAGCTGGGG 451
Db 338 GAGACCTGTGAGAGGATGCGAGTCTGCGCTGAGAGACCTTCAAGCTGGGG 397
Qy 452 AGCTTATCATAGAGAGACATATCTGCAAGTGTCAAGAGAGGCTAGTCTTTG 511

Db 398 AGCGACACACCGGAGAGAGCANNCTGTCCGGGGCCCGAGAGCCGACTCATAGGCTG 457
Qy 512 GAGTATCGTCTCTCAACAGCACTGATGACAGCTGTTCATGCGACGAGTGATG 571
Db 458 GTNNNNNNNGAGAGAGGCGCCAGGGGCCAGCAAGCTGTTGTGGCACTGCTGTTNNNN 517
Qy 572 GGAAGCCCGAGTATTTTCCACCATCTCCAGCGGAAACTGACCAAGACTGAGCGG 631
Db 518 NNN 577
Qy 632 ATGCGATGTTGCGGATGCTCTTCATGATGAGTGTGCTCGATGATTAAGCTCTT 691
Db 578 NNN 637
Qy 692 CGAGACCTTACCATCATCTCCGACTTGGATATCTACTATGCTATGTTTAGAGTG 751
Db 638 CAGACAGCTGCTCTTGTACCCCTGCTTGGACATCTACATCACTGAGCTTGTGAGG 697
Qy 752 GCACTTGTCTACTTTTGAACCTTCAACCTGAGATGCTGTCTCCAGAGCTCCAGCA 811
Db 698 CTCTCTTGTGATCTTCTGACGCTGACGCTGACGACCCAGAGAGAGCTGAGACAG 757
Qy 812 CCAAGAGAGAGTGTATACATCCAGCTGTGAGCTTTGCAAGAGAGACAGCTTGA 871
Db 758 CGGCGGAAATTTTACGTCACAGATGCTGCGCATGTGCGGGAGACTNNAGTTCT 817
Qy 872 ACTCTATGATGAGAGGAGCCCATTTGCTGTAGAGCGAGGAGTACCCGCTGCTG 931
Db 818 ACTCATACGTGATTCATTCCTCATGAGCTGCTGCTGCGGCGGATGAGTACCGTGTG 877
Qy 932 AGCTGCTTACCTGTCCAAAGCGGGGCGCTGCTTGGAGAGACCTTGGATTCATCAG 991
Db 878 AGAGCGCCACTGCGCAAGCTGCGCTGCTGCTGCGGAGGCTGCGGCTGCGGCTG 937
Qy 992 ATGATGACCTGCTCTTCAACCTTTCACAGGCGGAGAGCGGAAATATTCCTG 1051
Db 938 ATGAGACATCTTCTTCAATCTTCTTCAAGGCGGAGAGAGAGCGGCGGCGGCGG 997
Qy 1052 ATGAGTGGGCTGCTGATCTTCAATCTTGAAGCAATATGACCGCATTAAGAGCGG 1111
Db 998 GGAAGACATCTCTGCTCTTCACTTCACTTCACTTCACTTCACTTCACTTCACTT 1057
Qy 1112 TGCACTTGTATCGGGGAGAGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1171
Db 1058 TCAAGTCTGCTATCGGGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1117
Qy 1172 TCCCTGACAGAGTGCCTCTTAAACATGACATTAATCTTGTGAGCTTGAATGATG 1231
Db 1118 TGCCCTGATCAACACCCCATGAGATCAACGCAACTTCTGTGGCTGTGTGAAC 1177
Qy 1232 CTCCTGAGAGTGTCCGATGAGTGTGTGGAATTCCTGCTTCAAGAGAGAGAGAG 1291
Db 1178 AGCTCTGAGAGGCTGATGATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1237
Qy 1292 GATGAGCTGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1351
Db 1238 GATGAGAGAGTGTGCGCTTACATCGGCGAGAGAGAGAGAGAGAGAGAGAGAGAG 1297
Qy 1352 AAAGTGAAGCTGAGAGAGATCGGGGTGATGAGAGAGAGAGAGAGAGAGAGAGAG 1411
Db 1298 GCAAGCGCAGNN 1357
Qy 1412 AGAGGTGAGGTGTGAGACCCGCGCCAGTCTCCGAGATATGCGCTTCCAGAGACC 1471
Db 1358 NNN 1417
Qy 1472 ACGAGCAATCTTACATGATGATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1531
Db 1418 NNN 1477
Qy 1532 GTGAGTATCAGAGTGTGGGAGAGTGTGTGCTGAGAGAGAGAGAGAGAGAGAGAG 1591

[illegible]

QY	2669	TCOCCTCCCATGTCACAGGTCGTGGGTGGAGTGGAGCCCTTTAGTGAATGGTATCATCC	2728
Db	2615	TGGGC-----CTGGGGGTGGCTGGCGTGGCTTGCACTCCATTCGAGCCGAGTACATCA	2668
QY	2729	CTGCAGAAACGATCGTGTGTGTGAGATGGGGAGGC--CAAGCCCAACGACATGCAAGCT	2785
Db	2669	GTGCTNNNNNNNATCGTGTGTGATGGAGGAGTGCCTGGTGGCCAGCCCGCCGGGGGC	2728
QY	2786	TCGTGGAGATCTGGGTGGCTGTGTGTGGCTGTGAATCATGGCCCGGTCTCTCAAGCTCT	2845
Db	2729	CCGTGAGCTGTGTGTGGGTGACTGNNCAGCCGACTTCGGCAACGCACTCGAGACANNNN	2788
QY	2846	ATTACTTCATGACACTGACTCTCTCAGATCTGAAGCCCAAGCCGGGGGCCCATGTCCGGAG	2905
Db	2789	NNNNNNNNNNNACCCCAACCTTTGACCAATGTGTCCCAACCGTGGCCCGCGTCCGGG	2848
QY	2906	GCACCCAGTGCACATCAAGGCAAGGCACCAACTGAATGCGGAAGAAAGTGGTGTGATGT	2965
Db	2849	GCACACGGCTTACATCTCAGGACAGCTCTGATGTGGCAGAGAGGTCACATGACTG	2908
QY	2966	TTGGAAGCAGCCCTGTCTCTTCCACAGCGATCTCCATCTCATTTGTCTGCAACACCA	3025
Db	2909	TGAGGGACAGCGAGTGCAGTTGTGTAAAGAGATGCAAGCGATGTTGTGATCTCAC	2968
QY	3026	CATCCTCAGATGAGAGTCTG--AAATGAAGGTGTGGTGCAGTGGAGTGAACAGGGCCAGA	3082
Db	2969	CTGTCTCAACCTTGGGCCCAAGCCAGGCCCCCATTCACATTGCCATTGACCGGGCTAACA	3028
QY	3083	TCACACAGAACCTGGTCTTTCAG--TATGTGAAGACCCCAACCATGTGTGGATTGAGC	3139
Db	3029	TCTTCAAGCCCCGGGGTTCATCTACACTCACTCACTGAGACCCACCGTACCCGCTTGAGC	3088
QY	3140	CAGAAATGAGCATGTGATGAGAAACACACCATGCGCTGATGGGAGACCCACTGAGAC	3199
Db	3089	CCACTTGGAGCATATCA--NN	3145
QY	3200	TCATACAGAAACCCCAAGATCCGTGCAAGCATGAGGGAAGAGACATCAATATCTGTG	3259
Db	3146	NN	3205
QY	3260	AGGTTCTGAACGCTACTGATGACTGTGAGCGCCCGCCCTGCTCTGGGTCTTGAC	3319
Db	3206	NN	3265
QY	3320	ACCACTCAGACCTGACCGAGAGGCCCAAGAGTTTGGCTTCATCTTGACAACGTCCAGT	3379
Db	3366	NN	3325
QY	3380	CCCTGCTCATCTCAACAGACCAACTTCACTTACCAACCCGGTGTGTAAGGCT	3439
Db	3326	NN	3385
QY	3440	TTGGTCCCTCAGGAATCCTGGAAGCTCAAGCTGCAAGCCCATCATCTTAAGGACAGA	3499
Db	3386	NN	3445
QY	3500	ACCTGATCCGCTGTGAGTGGGGCAACGTGAAGTGAATACTAGTGTGTTGGG	3559
Db	3446	NN	3505
QY	3560	AGAAAGCGTGACCGTGACCGTGTCAATGTCCAGCTGCTTGGAAGTCCCCCAACCTCA	3619
Db	3506	NN	3565
QY	3620	TCGGCAGAGCAAAAGTATGGCCGTGTCCGTGATGAGTACTCCCGGGATGCTGT	3679
Db	3566	NN	3625
QY	3680	ACATTGCCCGGACAGCCCGCTAGCTGCCCGGCATCTGCAGATGCAATGCAATGGCTGGCG	3733
Db	3626	ACATCTCGGACAGACGGGCGCTGANNNNNNCCGGCAATGATGGGCTGGCGGCGGGGGCG	3685

Db
3626 ACATCTCGGCAAGCGGGCGCTGANNNNNNCCGGGCATGATGGGCGCTGGCCGGCGGGGGCG 3685

QY 3740 GCCTTCATCATTTTCATCGTGGCCGTGCTCATTTGCTATAAGCGCAAGTCCCGGAAA 3799
 Db 3686 GGGCTCTCTGCTGGCCATCAAGGCCGTGTGTGGCTTCAAGGGCAAGACNNNNNAAG 3745
 QY 3800 GTGACCTCAAGCTGAAGCGGCTGCGATGCAATGGAACAACCTGAGTCCCGTGTGCCC 3859
 Db 3746 CGGACCGTACNNNAAGCGCTGTGACGCTNNNNATGGAACAACCTGAGT----- 3793
 QY 3860 TGAGTGAAGGAAGCTTTGTGCGAGTGCAGACGGAATCATGAGCTGACAGTGAAC 3919
 Db 3794 NNN 3853
 QY 3920 TGGATGAGCGGGATTCCTGTTCTGGAATAGAACTTACACATGCGGGTGTCTTCC 3979
 Db 3854 NNN 3913
 QY 3980 CAGGAATTGAAGACACACCTGTCTCCGGAACCTTGAAGTCCCGGCTACCGGAGAGC 4039
 Db 3914 NNN 3973
 QY 4040 GTGTGAGAAAGCGCTGAAGCTTTGCCAGCTCATCAACAAGGTGTTCTGCTGT 4099
 Db 3974 NNN 4033
 QY 4100 CCTTCATCCGACGCTTGAAGTCCGACGTAAGCTTCCATGCGGAGCCGTGGAACGTG 4159
 Db 4034 NNN 4093
 QY 4160 CCTCATCATCAACGCGTCTGCGAGCAAGCTGAGTGAAGCCCATGATGTGTGAAGC 4219
 Db 4094 NNN 4153
 QY 4220 AGCTGTGCGGACCTCATTTGACAAAGACCTGAGAGCAAGCAACCTTAAGCTGCTGC 4279
 Db 4154 NNN 4213
 QY 4280 TCAGAGAGACTGAGTCAAGTGTGAGAGATGCTGACCAATGTTACTTCTCTCTCT 4339
 Db 4214 NNN 4273
 QY 4340 ACAAGTTCCTCAAGAGTGTGCTGGGAGACCTCTCTCTCTCTCTCTCTCTCTCTCT 4399
 Db 4274 ATTAAGTTCTGAAGAGTGTGCTGGGAGACCTCTCTCTCTCTCTCTCTCTCTCTCT 4333
 QY 4400 AGAGATGAGAGAGGCGCCCATTTGACGCGCATACGCGGAGCGGCTACTCTTGAAGC 4459
 Db 4334 AGCAGATGAGAGAGGCGCCCATTTGATGATGATGATGATGATGATGATGATGATGAT 4393
 QY 4460 AGGACAGGCTCATCGCGACAGAGATTTGACTAATAAACCCTGTCTGTGAGTGTGACCC 4519
 Db 4394 AGGACAGGCTCATCGCGACAGAGATTTGACTAATAAAGCACTGACCTGACCTGCTGTC 4453
 QY 4520 CAGACATGCGCAACAGCCCGGAGGTCCAGTAAGATCCTCACTGTGACACATCACTC 4579
 Db 4454 CGGAGACAGAGGCGGAGCGCCAGGTCCAGTGAAGTTCTCACTGTGACACATCACTC 4513
 QY 4580 AGGTCAAGAGAGATTTGATGATGATGATGATGATGATGATGATGATGATGATGATG 4639
 Db 4514 AGGCGCAAGATTAAGTGTGTGACATGTGTACAAAGGCACTTCGTAATCCAGCTCCCA 4573
 QY 4640 AAGCTGACATATGATTTGATGATGATGATGATGATGATGATGATGATGATGATGATG 4699
 Db 4574 AAGCTGACATATGATTTGATGATGATGATGATGATGATGATGATGATGATGATGATG 4633
 QY 4700 ATGAACAATCATCACCAAGATTTGATGATGATGATGATGATGATGATGATGATGATGATG 4759
 Db 4634 ATGAAGATGTCACACCAAGATTTGATGATGATGATGATGATGATGATGATGATGATGATG 4693
 QY 4760 ACCAGGTGCGAGATGTTCCGTGTGTGATGATGATGATGATGATGATGATGATGATGATG 4819
 Db 4694 ACCAGGTGCGAGATGTTCCGTGTGTGATGATGATGATGATGATGATGATGATGATGATG 4753
 QY 4820 CAGTGAACAATCATCACCGTCTCAGAGACCTCAGCAAGTAAATATGAAAAATGATCCGCT 4879

Db 4754 TGGCCAACTCTTCACTTACCTTACCGG---CTCCCTGAGCGGCTACGNNNNNNNNNNNN 4810
 QY 4880 ACACGGGCAAGCCCGGACAGCTCCGCTCAAGCACTATGATCACTCTGATCTGAGAA 4939
 Db 4811 NNN 4870
 QY 4940 GTGAGTCAAGATGTGCACTTATGAGAAACAAGAGACGGAAGACGGAAGAGGGGG 4999
 Db 4871 NNNGCAACNNATTTGTGCACTGTGTNNNNNNCACGACATGCCAATGCCAGAGGGG 4930
 QY 5000 ACCGGGGAGCAAGATGTGTGAAATCTACCTGACCCGACTCTCGGCACTAAGGGA 5059
 Db 4931 ACCGTGCGAGNNNNNTGTCTCANNAGATCTACTGNNNNNNNNNNNNNNNNNNNN 4990
 QY 5060 CACTGCAAGATTTGTGATGACCTCTTTGAACAATCTTCAAGACGGCAACCGTGT 5119
 Db 4991 CACTGCAAGATTTGTGATGACCTCTTTGAACAATCTTCAAGACGGCAACCGTGT 5050
 QY 5120 CTGCGCTGCGCCCTGGGCACTCAAGTATGATGATGATGATGATGATGATGATGATGATG 5179
 Db 5051 CGGCGCTGCGCCCTGGGCACTCAAGTATGATGATGATGATGATGATGATGATGATGATG 5110
 QY 5180 ATGCAATTCATGACCCGCAAGTCCGCAATACCTGGAAGAGCAATTTGCTGCTGAGT 5239
 Db 5111 GCCAGATCAGCGACCCCGACGTGCGCCACACTGGAAGAGCAATTTGCTGCTGAGT 5170
 QY 5240 TTTGGGTCAATGATCAAGAAACCCGCAATTTGTTGATGATGATGATGATGATGATGATG 5299
 Db 5171 TCTGGGTGATGATCAAGAAACCCGCAATTTGTTGATGATGATGATGATGATGATGATG 5230
 QY 5300 CAGACGCTGCTCTCTGTGTGTGCTCAGACCTTCAAGTACTTCTCTCAAGTCAAGC 5359
 Db 5231 CGGATGCTGCTCTGTGTGTGTGCTCAGACCTTCAAGTACTTCTCTCAAGTCAAGC 5290
 QY 5360 ACCGGCTGGGCAAGACTGCGCTCCCAAGTGTGTGATGATGATGATGATGATGATGATG 5419
 Db 5291 ACCGGCTGGGCAAGACTGCGCTCCCAAGTGTGTGATGATGATGATGATGATGATGATG 5350
 QY 5420 ACAAGATTTGGTGAAGAGTATTACTCAGACATAGGAAATGCCAGCATCAGGAGC 5479
 Db 5351 ACAAGATTTGGTGAAGAGTATTACTCAGACATAGGAAATGCCAGCATCAGGAGC 5410
 QY 5480 AAGACATGAAGCAATCTGTGCTGAGCAAGTCCGGAATGCAATGAATGAATTAACACA 5539
 Db 5411 AGGACATGAATGCTTACTGTGTGAGCAAGTCCGCTCCACGCGACGACTTGAAGCTC 5470
 QY 5540 TGAATGACCTCAGAGATCTTCTCTATTTGAGCAAAATCAGCGAGAG 5589
 Db 5471 TGAATGACCTCAGAGATCTTCTCTATTTGAGCAAAATCAGCGAGAG 5520

RESULT 7
 AY420220 5341 bp DNA linear GSS 12-DEC-2003
 LOCUS
 DEFINITION Pan troglodytes HCM7151 gene, VIRAL TRANSCRIPT, partial sequence,
 genomic survey sequence.
 ACCESSION AY420220
 VERSION AY420220.1 GI:39776177
 KEYWORDS
 SOURCE
 ORGANISM
 Pan troglodytes (chimpanzee)
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Pan.
 REFERENCE
 1 (bases 1 to 5341)
 Clark,A.G., Glanowski,S., Nielson,R., Thomas,P., Kejarival,A.,
 Todd,M.A., Tanenbaum,D.M., Clavello,D.R., Lu,F., Murphy,B.,
 Ferriter,S., Wang,G., Zheng,X.H., White,T.J., Slatney,D.J.,
 Adams,M.D. and Carcilli,M.
 Inferring nonneutral evolution from human-chimp-mouse orthologous
 gene trios
 JOURNAL Science 302 (5652), 1960-1963 (2003)
 PUBMED 14671302

QY 2196 TGGGCAAGCTGCTACGATGATCTCTCAATTCAGGGCAGCAGCAGCTGCCGC 2255
 DB 1861 AGGCAAGCTGATATGAGTGCCTCTCCACATCCCGGGCAGCCCGCGCTGCACCC 1920
 QY 2256 CCGCGCTTCAACAGCTCCAGGCTAGACATGCGCAAAACCTC--TTATTCATGAAG 2312
 DB 1921 CTTGCGCTTCAACACTCCAGCTCAGTGCAGAAATTCCTGANNNNNNNNNNNNNN 1980
 QY 2313 GATGAGATCAACAACCTGCCGTGAGTTGACAGTGTGTGGAATGGGCACTTCAAT 2372
 DB 1981 NNN 2040
 QY 2373 TGACACCCAGCTCAGAAATAAGTTCACTCTCAAGTGTGAGCCATGCTGAGAGCTG 2432
 DB 2041 NNN 2100
 QY 2433 CGGCTGTGCTCAAGGCTGACCCAGACTTCCGATGTGCTGCTGCAGGCGCCAGGCA 2492
 DB 2101 NNN 2160
 QY 2493 GTGCACCTTGCGCAGACTGCCCTGCCAGAGAGCAGTGCCTGAGAGTGTGTGCTG 2552
 DB 2161 NNN 2220
 QY 2553 CAAAAGCAGTGCACAAAACCCCGCATCACAGATAATCCGATGACAGGCCCGGGA 2612
 DB 2221 NNN 2280
 QY 2613 AGGGGAGCACAAGTCACTATCCGAGGGAGAACTGGGCGTGAATTTCGCGCATCCG 2672
 DB 2281 NNN 2340
 QY 2673 CTCCTCAGTCAAGTGTGCTGCGGTGAGTGCAGCCCTTAGTGGATGATCCCTGC 2732
 DB 2341 NNN 2400
 QY 2733 AGAAGCAGTGTGTGAGATGGGGGAGGCCAAGCCC--AGCCAGCATGCAAGCTTCT 2789
 DB 2401 NNN 2460
 QY 2790 GGAATCTGCGTGTGTGTGTGCGCTGATTTATGCCCCGATCTCAAGCTTAATTA 2849
 DB 2461 NNN 2520
 QY 2850 CTTATGACACTGACTCTCTCAGATGTGAAGCCAGCGGGGCCCCATGTCCGAGGAGAC 2909
 DB 2521 NNN 2580
 QY 2910 CCAAGTACATCAACAGCACAACCTTGAATGCCGGAAGCAAGTGTGTGTGTGTG 2969
 DB 2581 CTGATTTGGCATGAGGAGCACTGGAAGCAGGAGTGTGTGTGTGTGTGTGTGTG 2640
 QY 2970 AAAGAGCCCTGTCTCTTCCAGAGCGATCTCACTCATTTGTGCAACCAATC 3029
 DB 2641 TGGCGGCGCTGTCTCTTCTCTGNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN 2700
 QY 3030 CTCAGATGAGTGTAGATGAAAGTGTGCTGAGGTGAGGAGGCGCAAGATCCACCA 3089
 DB 2701 NNN 2760
 QY 3090 GGAAGTGTCTTTGATGTGGAAGACCCACATGTGTGAGTTAGGCAAGATGAG 3149
 DB 2761 NNN 2820
 QY 3150 CATGTCAAGTGAACAACACCATGCGGTATGGGGAGCCCACTGACCTCATACAGAA 3209
 DB 2821 NNN 2880
 QY 3210 CCCCCAGATCCGTGCAAGCATGAGGAGAGAGACATCAATATCTGTGAGTTCTGAA 3269
 DB 2881 NNN 2940

QY 3270 CGTACTGAGATGACCTGTGACGCGCCGCGCTGTGAGGTCTGACCA---CAATG 3326
 DB 2941 NNN 3000
 QY 3327 AGACCTGACCGAGAGGCGCGAGAGATTGCTTATCTTGAACAACCTCAATCTCTGCT 3386
 DB 3001 AGAGCTGGGAGAGCGCGAGTGAAGTGTGCTGTGATGACAAATGTGCGCTCTGCT 3060
 QY 3387 CATCTGAAACAAGCAACCTTCACTCTATCCCAACCGGTGTTGAGGCTTTGATCC 3446
 DB 3061 TGTGCTCACTCACT 3120
 QY 3447 CTCAGAAATCTGAGTCAAGCTGACCGCCATCATCTTAAGGCGAAACCTGAT 3506
 DB 3121 CACTGGCTGTGAGGTGAAGCCAGCTCCCACTCATCTCAAGGCGCGAACTCTT 3180
 QY 3507 CCGGCTGTGCTGGGCGCACTGAAAGCTGAATCACTGCTGTGTTGGGAGAGCC 3566
 DB 3181 GCCACTGCACT--GGCACTCCGACTCACTCACTCACTGCTCATCTGCTCAAC 3237
 QY 3567 GTGACCGTGAACGCTGATGATGTCAGCTGCTGAGAGTCCCAACCTCATCGGAG 3626
 DB 3238 CTGTACCTCACTGTGTGAGAGAGCACTGTGTGAGAGGCGCCAACTCATCGGANN 3297
 QY 3627 GCAAAAGTGAATGGCGGTGTGCTGATGAGTATCTCCCGGGAGTGTGTATGCTG 3686
 DB 3298 NCACAAAGTCAACGCTGCGGAGAGTGTGCTGAGTTCTGCGCAAGGACACTGAGTGA 3357
 QY 3687 CCGGAGAGCGCGCTGAGCTGCGGCGCACTGTCAGATGCACTGCTGCGGCTCTCT 3746
 DB 3358 CTCGACAGCTGCTGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 3417
 QY 3747 CATCATTTTCATGCTGCGCTGTCTCATTTGCTATTAACGCAAGTCCCGAAATGACT 3806
 DB 3418 GCTGCTGTATCGGT 3477
 QY 3807 CACGCTGAAGGCGCTGACAGATGAGATGAGCAACCTGAGTCCGCTGAGCTGAGT 3866
 DB 3478 CACACTCAAGCGGCTGCACTGCACTGCACTGCACTGCACTGCACTGCACTGCA 3537
 QY 3867 CAAGAGCCTTTGCGAGCTGCAAGCAGCATCATGATGAGTCACTGATGAGT 3926
 DB 3538 CAANN 3597
 QY 3927 AGCGGAGTCCGTTCTCTGACTATGAATTAACCACTGAGGCTGCTCCAGAGAT 3986
 DB 3598 TGGCGGATCCCTCTCTGACTACCGGACATGAGCACTGAGGCTCTTCTCTGAGAT 3657
 QY 3987 TGAAGACCACTGTCTCCGAGACTTGAAGTCCCGGCTTACCGGAGAGCGTGTGA 4046
 DB 3658 CGAGGACCACTGTGTCTCAAGAGATGAG-----NNNNNNNNNNNNNNNNNN 3705
 QY 4047 GAAAGCTGAAAGCTCTTGTGCGGAGCTCATCAACAAGGTGTTCTGTCTCTAT 4106
 DB 3706 NNN 3765
 QY 4107 CCGCAGCTGTAGTCCAGCGTAGTCTCTCATGCGGACCGTGGCAAGTGTGCTCACT 4166
 DB 3766 NNN 3825
 QY 4167 CATCATGACCGTGTGCAAGCAGCACTGAGTGCCTGAGTGTGAGAGAGCTGCT 4226
 DB 3826 NNN 3885
 QY 4227 GGGCGACCTCATTTGACAGAACTTGAGAGCAAGCAACCACTTACTGCTGAGAG 4286
 DB 3886 NNN 3945
 QY 4287 GACTGAGTCAAGTGTGAGAGATGCTGACCAATGTTTACTTCTCTCTTACAAGT 4346
 DB 3946 GACTGAGTCCGTGGAGAGAGATGCTAACTAATGCTTCACTTCTCTGTATAGTT 4405
 QY 4347 CCTCAAGAGTGTGCTGGGAGCCCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 4406

Db	4006	CCTCAGGAAGTGGCGCTGGGGAGCCGCTGTTCACTGCTACTGGCCATCAAGACGAGAT	4065
QY	4407	GGAGAGGGGCCCATTTGACGCATCAACGGCGAGGCCCGCTACTCTTTAGCGAGACAA	4466
Db	4066	GGAGAAAGGGCCCATTTACCGCATCAAGGGGTGAAGCAGCTACTCCTGTAGTGAACAA	4125
QY	4467	GCTCATCCGCGACGACGATTTGACTACAAACCCCTGGTCTTAGCTGTGTCAAGCCAGACAA	4526
Db	4126	GCTCATCCGAGGACGATTTGACTACAAACNNNNACCTGAATGTGTGAACCTGTGAAA	4185
QY	4527	TGCCAACACCCCGAGTCCCAATAAAGATCTCAATGTGACACCATCATCAAGTCA	4586
Db	4186	TGAGAAATGCACTTAGGTCGCGGTGAAGGGCTGGAATGTGACACGGTACCCAGGCAA	4245
QY	4587	GGAGAAATTTGGATGTCATCTTCAAGAAATGTCTTGCTCCACCGGCGCAAGAGTCG	4646
Db	4246	GGAAGATGTTGGACGCTGCTTCAAGAGGGGTGCCNNNNNNNNNNNNNNNNNNNNNN	4305
QY	4647	AGATATGATCTGAGTGGCGACAGAAAGTGGGGCAAGATGATCTTGACAGATGAGA	4706
Db	4306	NN	4365
QY	4707	CATCACCAACAAGATTGAATGATTTGGAAGCGACTGAACAACCTGGCCACTACAGAT	4766
Db	4366	NN	4425
QY	4767	GCAGATGTTCCGTGTGTGCGATTAGTGTCCAGCAGGTGACAGCTTAACGACGTGA	4826
Db	4426	GACAGAAGGGTCTCGGGGCACTGGTGCCCAACAGACATCCGCTACATATCTTCCAA	4485
QY	4827	CAATCCACCGCTTCCAGGACCTCAGCAAGTAAATAGAAAACATGATCCGCTACACGGG	4886
Db	4486	CTCTCCACCTTCAACAAATCCCTTC--AGCAGATACGAGACATGCTGGCGACGGCAG	4542
QY	4887	CAGCCCCGACAGCCTCGGCTACGCGACACCTATGATCATCTGACTGTGAAGTGAAT	4946
Db	4543	NNNNCCGACAGCCTCGGCTCGCGACCGCTATGATCACGCCACCTGGAGAGCGCAC	4602
QY	4947	CAAGATGGGACACCTAGTGAAGAACAGACACAGAGACAGAAAGGAGGGGACCGGGG	5006
Db	4603	CAAGGTGTGGCACCTGGTGAAGAACAGACACACCTGACACAGGTGAGGTGACCGCG	4662
QY	5007	GAGCAAGATGGTGTGTAATCTACCTGACCCGACTCTGGCCAATAAGGACACATGCA	5066
Db	4663	CAGCAAGATGGTCTCGAGATCTACTGACACGGCTACTGGCCACCAAGNNNNNNNNNN	4722
QY	5067	GAAATTTGTGATGACCTCTTGAGACATCTTGACACGGCACACCGTGGCTGTGCGCT	5126
Db	4723	NN	4782
QY	5127	GCCCTGGCCATCAAGTACATGTTGACTTCTCTGGATGAGACAGGCTGTAATAATGCAAT	5186
Db	4783	GCGCTGGCNATCANNNNCAATTTGCACTTCTGTGATGNNNNNNNNNNNNNNNNNNNN	4842
QY	5187	TCATATACCCGACGTCGCCCATACTGTGAAGACAAATGGCTGCCCTGAGGTTTTGGGT	5246
Db	4843	NN	4902
QY	5247	CAACATGATCAAGAACCCGACAGTTTGTGTTGACATCATAGAACAGCATCAACAGCGC	5306
Db	4903	GAACTGATCAAGAACCCACAGTTTGTGTTGACATCAAGAACAGCATCAACAGCGC	4962
QY	5307	CTGGCTCTGTGTGGTGGCTCACAAGCTTCAATGACCTTGCTCAAGTCAAGTCAAGCGCT	5366
Db	4963	CTGGCTCTGTGTGGTGGCCCAAGCTTCAAGNNNNNNNNNNNNNNNNNNNNNNNNNNNN	5022
QY	5367	GGGCAAGACATCGCCCTCCAAACAGCTGCTATATCCAAAGACATCCCAAGTACAGAA	5428
Db	5023	NN	5082
QY	5427	TTGGGTGGAGAGTATTACTCAGACATAGGAGAAAGTGCAGGCAATCAAGACAT	5486

[illegible]

Hayashida, K., Hayatsu, N., Hiramoto, K., Hiraoka, T., Hirozane, T., Hori, F., Imotani, K., Ishii, Y., Itoh, M., Kagawa, I., Kasukawa, T., Kato, H., Kawai, J., Kojima, Y., Kondo, S., Kono, H., Kouda, M., Koya, S., Kurihara, C., Matsuyama, T., Miyazaki, A., Murata, M., Nakamura, M., Nishi, K., Nomura, K., Numazaki, R., Ohno, M., Ohsato, N., Okazaki, Y., Saito, R., Saitoh, H., Sakai, C., Sakai, K., Sakazume, N., Sano, H., Sasaki, D., Shibata, K., Shinagawa, A., Shiraki, T., Sogabe, Y., Tagami, M., Tagawa, K., Takahashi, F., Takaku, Akahira, S., Takeda, Y., Tanaka, T., Tomaru, A., Toya, T., Yasunishi, A., Muramatsu, M. and Hayashizaki, Y.

Direct Submission
Submitted (16-JUL-2001) Yoshihide Hayashizaki, The Institute of Physical and Chemical Research (RIKEN), Laboratory for Genome Exploration Research Group, RIKEN Genomic Sciences Center (GSC), RIKEN Yokohama Institute; 1-7-22 Suenho-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan (E-mail: genome-res@sec.riken.go.jp, URL: http://genome.sec.riken.go.jp/, Tel: 81-45-503-9222, Fax: 81-45-503-9216)

CDNA library was prepared and sequenced in Mouse Genome Encyclopedia Project of Genome Exploration Research Group in Riken Genomic Sciences Center and Genome Science Laboratory in Riken Division of Experimental Animal Research in Riken contributed to prepare mouse tissues.

Please visit our web site for further details.
URL: http://genome.sec.riken.go.jp/
URL: http://fantom.gsc.riken.go.jp/
Location/Qualifiers
1. .3355
/organism="Mus musculus"
/mol_type="mRNA"
/strain="C57BL/6J"
/db_xref="PANTOM DB:C330023B09"
/db_xref="WGI:2416574"
/db_xref="taxon:10090"
/clone="C330023B09"
/cell_type="ES cells"
/clone_lib="RIKEN full-length enriched mouse cDNA library"
1. .2210
/note="unnamed protein product; PLEKXIN 3 homolog [Mus musculus] (SPTK1P70208, evidence: PASTY, 99.7%id, 39.2%length, match=2205)
putative"
/codon_start=3
/protein_id="BAC33681.1"
/db_xref="GI:26340036"
/translation="DYKPSGSHVYLKKNLIPAAAGSSRLNTVILGGPCALTVSDTO LLDPSBQTRQFVWLVGSLFELGTLHTADRALTPAMGLAAGGLLALATTV LVAYKRTQADRLKRLQOMDNLESRALECKEFAELQTDINLNMGVOLPE LDRVYAVRLVPEIGIEAHVPLKEIDTPVNEKALRLFGQLHSRAFLPIHLEAS SFGMRDGTVAALVVALOSGLDVALGKQLADLIEKLLSKNPKLILRLEBSA EKLLNMFPLFLKPLKEKAGEPLFLYCAIKOEMKEGIDATGEGARVLSBEDKLR QQIDYKTLTLHCICPESESAOVPVAVLNDSTLTKDKLIDATGEGARVLSBEDKLR MDLBMROGRPARILLODEDTITTECDMKVNSLAAYOTDSIVLAVKQVAYMA NSFTFRSLRYSLELPAASPSLSKRAPLTPDOBAKAMHLYRNDHPDHEGD RGSQWSEIYLRLATKGLKQFVDLFEVTSIHRGSAFLAKIKVFLDPEAD QKQISDPDVHTWKSNCPLRFWVNIKNQPFVDIHKNSIDACLSVAQVPMDSCT TSEHRLGDSFENKLYAKDIYNTKSWERYDIARMSISPDQDAVLVQSRLLHA NDPVLSALSELVYVTKYRGELITSLDRDASCRKHLRQKLEQIITLVSSSS"

polya_signal
3332..3337
/note="putative"
3355
polya_site
/note="putative"

ORIGIN
Query Match 21.5%; Score 1221.2; DB 11; Length 3355;
Best Local Similarity 72.8%; Pred. No. 1,4e-246;
Matches 1622; Conservative 0; Mismatches 588; Indels 18; Gaps 3;

Db 1 TAGATGCAACCTGGTTCAATGTTATGAAAGGCAAGAACTATATCCCTGTCGAC 60
QY 3458 TGAAGCTCAAGCTGGCAAGCCCATCAATCTAAAGGCAAGAACTATATCCCTGTCG 3517
Db 1 TAGATGCAACCTGGTTCAATGTTATGAAAGGCAAGAACTATATCCCTGTCGAC 60
QY 3518 CTGGGGCAAGCTGAAGCTGAACCTGCTGTTGGGGAAGAAAGCCGTCGACGTA 3577

Db 61 CTGGCAAGCTCCCGC---CTCACTACACAGTGTGTAATGAGAGACAGCAATGTCATTA 117
QY 3578 CCGGTGAGATGTCAGACTCTCTGCGAGTCCCAACCTCATCGCAGGCAAGAGTGA 3637
Db 118 CTGCTCAGATCTCAACTCTGTGTGATTCGCCAAGCAGACAGCGGCAAGCTGTA 177
QY 3638 TGGCCGTGTCGTGTCAGTGAAGTACTCCCGGGATGTGTACATTCGCCGAGAGCC 3697
Db 178 TGGTGTGTCGTGTCGTGTCGTGTCGTGTCGTGTCGTGTCGTGTCGTGTCGTGTC 237
QY 3698 CGCTCAGCCCGCCGCGCATGTCAGATCGCAGTGTGTCGTGTCGTGTCGTGTCGTGTC 3757
Db 238 CACTGACCTTACACAGCTATGTGTGGGCTTACAGAGGGGGCGGCTTGTCTGTGCGCA 297
QY 3758 TCGTGGCCGTGTCATTTGCTTATTAACGAAATCCCGGAAAGTACCTGACCTGAAGC 3817
Db 298 TCACTGTTGTGTCGTGTCGTGTCGTGTCGTGTCGTGTCGTGTCGTGTCGTGTCGTG 357
QY 3818 GCGTCAGATGTCAGATGTCAGATGTCAGATGTCAGATGTCAGATGTCAGATGTCAGAT 3877
Db 358 GCGTTCAGCTACAAATGAGACCTGAGTCTGTGTGTCGTGTCGTGTCGTGTCGTGTCGT 417
QY 3878 TTGCGAGCTGCAAGCCGACATCATGAGTGAACCACTGATGAGCCGGGATTC 3937
Db 418 TTGCTGAGCTGCAAGCTGATATGATGATGATGATGATGATGATGATGATGATGATG 477
QY 3938 CATTCTGAGCTATGAACTTACACATGCGGGTCTGTTCCAGAGATGAGACACAC 3997
Db 478 CATTCTGAGCTATGAACTTACACATGCGGGTCTGTTCCAGAGATGAGACACAC 537
QY 3998 CTGCTCTCCGGGACCTTGAAGTCCCGGGCTACCGGCAAGAGCGTGTGAGAAAGCCCTGA 4057
Db 538 CAGTGTCTGAGAGAGCTGATGATGATGATGATGATGATGATGATGATGATGATGATG 585
QY 4058 AGCTCTTGGCCAGCTCATCAACAAGAGTGTCTGCTGTCTGCTGTCTGCTGTCTGCTG 4117
Db 586 GTCTATTGAGCAAGTCTGTCGACAGCCGTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 645
QY 4118 AGTCCAGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 4177
Db 646 AGGCCAGAGTATGTTCTTCATGCTGATGCTGATGCTGATGCTGATGCTGATGCTG 705
QY 4178 TGTGTCAGAGGAGCTGAGTACCCACTGATGTGTGTAAGACAGCTGTGCGGACCTCA 4237
Db 706 CCTGTCAGAGGCGGCTTGAATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 765
QY 4238 TTGACAAAGACCTGAGAGGCAAGAACACCTAGAGTGTGCTGAGAGAGCTAGTCA 4297
Db 766 TAGAGAAAACCTGAGAGGCAAGAACACCTAGAGTGTGCTGAGAGAGCTAGTCA 825
QY 4298 TGGCTGAGAGATGTCGACCAATGTGTTACTTCTCTCTCTCAAGTTCTCAAGAGT 4357
Db 826 TGGCTGAGAGATGTCGACCAATGTGTTACTTCTCTCTCTCAAGTTCTCAAGAGT 885
QY 4358 GTGCTGGGAGACCT 4417
Db 886 GGGCCGGGAGACCT 945
QY 4418 CCATTGAGCGCATACGCGGCGGCGCTACTCTCTTGAAGGAGACAAAGCTCATCCGCC 4477
Db 946 CCATTGAGCGCATACGCGGCGGCGCTACTCTCTTGAAGGAGACAAAGCTCATCCGCC 1005
QY 4478 AGCAGATTGACTACAAAACCTGTGCTGAGCTGTGTGAGCCGAGAAATGCCAAGACC 4537
Db 1006 AGCAGATTGACTACAAAACCTGTGCTGAGCTGTGTGAGCCGAGAAATGCCAAGACC 1065
QY 4538 CCGAGTCCAGTAAAGATCTCACTGATGACACATCATCAAGTCAAGAGAAAGATTC 4597
Db 1066 CCGAGTCCAGTAAAGATCTCACTGATGATGATGATGATGATGATGATGATGATGATG 1125
QY 4598 TGAATGCATCTTCAAGATGTGCTCTCTCCACCGGCGCAAGAGCTGCAATATGATC 4657

Db 1166 TGGATCTGTGTACAAAGGATATTCTCACTACGCGCCCAAGCTGAGACATGGATT 1185
 Qy 4658 TCGATGGGCAAGAAAGTGGGCAAGATGATCTTTCAGAGTAAAGACATCAACCA 4717
 Db 1186 TGGATGGGCGGAGGCGCGCATGCGCCGATATCTCTCCAGATAGAGACATCACTACAA 1245
 Qy 4718 AGATTGAGATGATGGAAGGACATGAAACACTGGCCCACTACCAAGGTGCCAATGGTT 4777
 Db 1246 AGATTGAGTGAATCGGAAGAGGTCACCTCAATGGCCCACTACCAAGGTGCTAATGGTT 1305
 Qy 4778 CCGTGGTGCATTAAGTGTCCAGACAGTGAAGCTTATAGCAAGTGAACAACTCAACG 4837
 Db 1306 CTTTATGACACTGTGTGCCCAACAGATGTCTGCTATTAATAGGCCAACTCGTTACCT 1365
 Qy 4838 TCTCCAGACCTGACAGATTAATATGAAAATATATCCGTTACCGGGCAGCCCGACCA 4897
 Db 1366 TCACCCCGG---TCACCTTATGCTACGAGAGCTTCTCTCGTCCAGCAGCCCGATTA 1422
 Qy 4898 GCGTCGCTCAGGACACCTATGATCACTGACCTGAGAGTGAATCAAGATGTGAC 4957
 Db 1423 GCGTCGCTCCGAGACCTATGCTACGCGCTGACCAAGAGGCAAGTGCCTCAAGCTGTGC 1482
 Qy 4958 ACCTAGTGAAGAACCGAGACCGAGACCGAGAGAGGAGGAGCCGAGGAGCAAGATG 5017
 Db 1483 ACCTGGTGAAGAACCGAGACCACTGATCACTGAGAGAGAGAGCCGCGAGCAAGATG 1542
 Qy 5018 TGTCTGAATCTAACCTGACCGGACTCTGCGCCCACTAAGGGGCACTGAGAGAGTTGTG 5077
 Db 1543 TGTCAAAATATATCTCAAGAGGCTGCTGCGCCCAAGGGGCACTGAGAGAGTTGTG 1602
 Qy 5078 ATGACCTCTTTGAGACCATTTTCAAGACGCGCAACCGTGTCTGCTGCTGCTGCGCC 5137
 Db 1603 ATGACTGTTTGAATCTGTGTGATGATGACCGCCAGCGGGGCTCAAGCTTACCTTGGG 1662
 Qy 5138 TCAAGTACATGTTTGAATCTTCTGATGAGACAGGCTGATTAATAGGCAATCAATGAC 5197
 Db 1663 TCAAGTACATGTTTGAATCTTCTGATGAGACAGGCTGATTAATAGGCAATCAATGAC 1722
 Qy 5198 ACCTCGGCATACCTGGAAGAGCAATGCTGCGCTGAGTGTGCTGCTGCTGCTGCTGCT 5257
 Db 1723 ATGTGGGTACACTGGAAGAGCACTGCTTACCTGCTGCTGCTGCTGCTGCTGCTGCT 1782
 Qy 5258 AGAACCGGCACTTGTGTGTGATGATCAATGAAGACATCAAGACGCTGCTGCTGCTG 5317
 Db 1783 AGAATCGCAATTTGTGTGTGATGATCAATGAAGACATCAAGACGCTGCTGCTGCTG 1842
 Qy 5318 TGTGGGTGAGCTTCTGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 5377
 Db 1843 TGTGGGTGAGCTTCTGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1902
 Qy 5378 GCGCTTCAACAGCTGCTGATGATGATGATGATGATGATGATGATGATGATGATGAT 5437
 Db 1903 CACCTTCAACAGCTGCTGATGATGATGATGATGATGATGATGATGATGATGATGAT 1962
 Qy 5438 GGTATTAATCAAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 5497
 Db 1963 GGTATTAATCAAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2022
 Qy 5498 TGTGGGTGAGCTTCTGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 5557
 Db 2023 TGTGGGTGAGCTTCTGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2082
 Qy 5558 TCTTCTCTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 5617
 Db 2083 TCTTCTCTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2142
 Qy 5618 AGTGTGGAAGCAAGAACTGGCTTCAAACTAAGACAGTCAATCAATCAATCAATCAAT 5677
 Db 2143 CTGTGTGGAAGCAAGAACTGGCTTCAAACTAAGACAGTCAATCAATCAATCAATCAAT 2202
 Qy 5678 AAGGCTGA 5685
 Db 2203 GCAAGCTGA 2210

RESULT 9
 CD653539
 LOCUS
 DEFINITION
 AGENCOURT 14542448 NIA Human H1 Embryonic Stem Cell cDNA Library
 (Long) Homo sapiens cDNA clone IMAGE:30420107 5', mRNA sequence.
 CD653539
 VERSION
 KEYWORDS
 SOURCE
 ORGANISM
 Homo sapiens (human)
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.
 1 (bases 1 to 838)
 NIH-MGC http://mgs.nci.nih.gov/
 National Institutes of Health, Mammalian Gene Collection (MGC)
 Unpublished (1999)
 Contact: Daniela S. Gerhard, Ph.D.
 Office of Cancer Genomics
 National Cancer Institute / NIH
 Bldg. 31 Rm10A07 Bethesda, MD 20892
 Email: gsabbs@mail.nih.gov
 Tissue Procurement: Irene Gintis and Mahendra Rao, NIA
 cDNA Library Preparation: Yulan Piao and Minoru Ko
 cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LINT)
 DNA Sequencing by: Agencourt Bioscience Corporation
 Clone distribution: MGC c
 can be found through the I.M.A.G.E. Consortium/LINT at:
 http://image.llnl.gov
 Plate: NDMA95 row: m column: 12
 High quality sequence stop: 650.
 Location/Qualifiers
 1..838
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /clone="IMAGE:30420107"
 /tissue_type="Embryonic Stem cells"
 /cell_line="WA01"
 /lab_host="DH10B (TI phage-resistant)"
 /clone_idb="NIA Human H1 Embryonic Stem Cell cDNA Library (Long)"
 /note="Vector: PCMV-Sport6; Site 1: NotI; Site 2: SalI;
 This is a long-transcript enriched cDNA library (Genome
 Res. 11: 1553-1558 (2001) [PMID: 11544199]) from WA01/
 cell line. Undifferentiated human ES cell line WA01/H1
 was obtained from WiCell Research Institute, Inc.,
 Madison, WI, cultured according to their instructions, on
 MEV feeders. They formed round colonies with defined edges
 and were positive for alkaline phosphatase, SSEA-4, OCT3,
 OCT4, REX1, UTR, TERT, SOX2, CX43 and CX45. They are
 negative for GAP42, GAP43, PDX1, NCM, MSX1, FLK1, SSEA-1,
 TUBB3, NES, GFAP, and BOWS. When confluent (18-10 days
 after plating), the ES cells from 4 x 6cm dishes were
 treated with 1 mg/ml collagenase, type IV
 (Invitrogen/GIBCO) for 5-10 min and gently scraped off
 with 5 ml pipette. RNA was purified with TRIzol Reagent
 from Invitrogen. Protocol ref: Genome Res. 11: 1553-1558
 (2001) [PMID:11544199] Double-stranded cDNAs were
 synthesized with an oligo(dT) primer [Invitrogen:
 5'-pGATAGTCTTAAATCCGAGCGCGCCCTTTTCTTTTCTTTT-3'] from
 3.49 of total RNA, treated with 14 DNA polymerase, and
 purified by ethanol-precipitation. The cDNAs were ligated
 to lone-linker Lp-SalI, purified by phenol/chloroform
 extraction, and separated from free linkers by
 Centricon-100 column. Then, the cDNAs were amplified by
 long-range high fidelity PCR using Ex Tag polymerase
 (Takara) with a primer SalI-S for 25 cycles. The products
 were purified by phenol/chloroform extraction and
 Centricon-100 column. The cDNAs were digested with SalI
 and NotI enzymes and cloned into SalI/NotI site of
 PCMV-Sport6 plasmid vector. The average insert size is

FEATURES

source

ORIGIN about 3.6kb. "

ORGANISM	Homo sapiens Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homidae. Homo.
----------	---

Query Match	13.1%	Score 743.8;	DB 14;	Length 838;
Best Local Similarity	96.4%	Pred. No. 4.9e-146;		
Matches 770; Conservative	0;	Mismatches 26;	Indels 3;	Gaps 1;

REFERENCE
1 (bases 1 to 1059)
AUTHORS
NIH-MGC <http://mgc.nci.nih.gov/>.
TITLE
National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL
Unpublished (1999)

3819 GCTGCAGATGCAGATGGAACAACCTGGAGTCCCGGTGTGCGCCCTGGAGTGCAGAGAAACCTT 3876
72 GCTGCAGATGCAGATGGAACAACCTGGAGTCCCGGTGTGCGCCCTGGAGTGCAGAGAAACCTT 131
3879 TGCCTGAGCTGCAGACGGACATTCATAGAGTGCACAGTGAACCTGGATGGAAGCCGGGAATTCC 3933
132 TGCCTGAGCTGCAGACGGACATTCATAGAGTGCACAGTGAACCTGGATGGAAGCCGGGAATTCC 191

FEATURES

Source

Л. 1059
6507. : Т

```

/ mol type = "mrna"

```

```
/db_xref="taxon:9606"
```

```
/clone="IMAGE:5/243/4"
/job host="PH10B"
```

```
/clone lib="NTH MGC 125"
```

```
/note="Organ: ovary (po
```

Site 1: EcoRV (destroyed), Site 2: NotI. RNA source pool of three ovaries, from females ranging in age from 38 to 49 ov. Library is λ gt10-dpn primed and directionally cloned (EcoRV site is destroyed upon cloning). Average insert size 2.1 kb, insert size range 1-3.5 kb. Library is normalized and enriched for full-length clones and was constructed by C. Gruber (Invitrogen). Research Genetics cranking code "C35."

ORIGIN

432 GCTGCAAGCAGCTGGAGTACGCACTGATGTGCTTGAGCAGCTGCTGGCCGACTCAT 491
4239 TGAACAAGACCTGGAGGCAAGAACCACTTAAGCTGTGCTTAGAGAGATGAGTCACT 429

Query Match	11.6%	Score 658.8	DB 12	Length 1059
Best Local Similarity	81.4%	Pred. No. 5.2e-128		
Matches 788	Conservative	0	Mismatches 177	Indels 3
				Gaps 2

RESULT	10			
LOCUS	BMS47417			
DEFINITION	BMS47417	1059 bp	mRNA	linear
AGENCY	CURR_6507625	NIH_MGC_125	Homo sapiens	cDNA clone IMAGE:572437
ACCESSION	BMS47417			
VERSION	BMS47417.1	GI:18781167		
KEYWORDS	EST.			
SOURCE	Homo sapiens (human)			

QY 4242 CAGAACCTGGAGAGCAAGAACACCCCTAAGCTGCTGCTCAGAGAGACTGATGATGATG 4301
 Db 491 TAAGAACTGGAGAGCAAGAACACCCCTAAGCTGCTGCTCAGAGAGACTGATGATGATG 550
 QY 4302 TGAGAGAGTGTGAGCAAGTGTGATGCTTCTCTCTCAAGTTCCTCAAGAGATGATG 4361
 Db 551 TGAAGAGATGTGAGCAAGTGTGATGCTTCTCTCTCAAGTTCCTCAAGAGATGATG 610
 QY 4362 TGGGAGAGCCCT 4421
 Db 611 AGGGAGAGCCCT 670
 QY 4422 TGAGCAGTACAG 4481
 Db 671 TGAATGCAATCAG 730
 QY 4482 GATTGACTACAAACCCCTGCTCTGAGCTGTGTGAGAGAGAGAGAGAGAGAGAGAG 4541
 Db 731 GATCGAGTACAAACCCCTGCTCTGAGCTGTGTGAGAGAGAGAGAGAGAGAGAGAG 790
 QY 4542 GGTCCCACTAAGAGATCTCTCACTGTGACACCTCACTCACTCACTCACTCACTCA 4601
 Db 791 GATCCCCAGTGAAGTGTAACTGTGACACCTCACTCACTCACTCACTCACTCACTCA 850
 QY 4602 TGGCATCTTCAAGATGT-GCCTGTGCTCCACCGGAGAGAGAGAGAGAGAGAGAGAG 4660
 Db 851 TGGCGGAGATGAAGATGTGGCCCTTATCCACAGGAGAGAGAGAGAGAGAGAGAGAG 910
 QY 4661 AGTGGGAGCAAGG--AAGTGGGAGCAAGATGATCTTGAGAGAGAGAGAGAGAGAG 4718
 Db 911 GATGGGAGCAAGGAG 970
 QY 4719 GATTGAGA 4726
 Db 971 GATTGAGA 978

RESULT 11
 AL529477 1201 bp. mRNA linear EST 23-MAY-2003
 LOCUS AL529477 Homo sapiens NEUROBLASTOMA COT 50-NORMALIZED Homo sapiens
 DEFINITION CDNA clone CS0DD006YR12 5-PRIME, mRNA sequence.
 ACCESSION AL529477
 VERSION AL529477.2 GI:31067320
 KEYWORDS EST.
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 REFERENCE 1. (bases 1 to 1201)
 Li, W.B., Gruber, C., Jessee, J. and Polayes, D.
 TITLE Full-length cDNA libraries and normalization
 JOURNAL Unpublished (2001)
 COMMENT On Feb 13, 2001 this sequence version replaced gi:12792970.
 Contact: Genoscope
 Genoscope - Centre National de Sequencage
 BP 191 91006 Evry cedex - France
 Email: seqref@genoscope.cns.fr; Web: www.genoscope.cns.fr
 was normalized. Library was constructed by Life Technologies, a
 division of Invitrogen. This sequence belongs to sequence cluster
 792.r For more information about this cluster, see
 http://www.genoscope.cns.fr/
 cgi-bin/cluster.cgi?seq=CS0DD006B060P1&cluster=792.r. Contact :
 Feng Liang Email: fliang@lifetech.com URL :
 http://fulllength.invitrogen.com/ Invitrogen Corporation 1600
 Faraday Avenue Genoscope Sequence ID : CS0DD006B060P1.
 Location/Qualifiers
 1. 1201

FEATURES

Source

/organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /clone="CS0DD006YR12"
 /issue_type="NEUROBLASTOMA COT 50-NORMALIZED"

ORIGIN
 /clone_lib="Homo sapiens NEUROBLASTOMA COT 50-NORMALIZED"
 /note="1st strand cDNA was primed with a NotI-oligo(dT)
 primer. Five prime end enriched, double-strand cDNA was
 digested with Not I and cloned into the Not I and EcoR V
 sites of the pCMVSPORT 6 vector. Library was normalized."

Query Match 11.4%; Score 650.6; DB 9; Length 1201;
 Best Local Similarity 82.5%; Pred. No. 3e-126;
 Matches 780; Conservative 1; Mismatches 160; Indels 4; Gaps 3;

QY 3773 TTGCTATTAAGCAAGATCCCGGAAAGTGAACCTCAAGCTGAAGCGGCTGCAGATGACA 3832
 Db 63 TTGCTATTAAGCAAGATCCCGGAAAGTGAACCTCAAGCTGAAGCGGCTGCAGATGACA 122
 QY 3833 TGGACAACTGAGATCCCGTGTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 3892
 Db 123 TGGACAACTGAGATCCCGTGTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 182
 QY 3893 CGGACATCAATGAGTGAACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 3952
 Db 183 CGGATATCAATGAGTGAACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 242
 QY 3953 GAACTTAACACATGCGGAGTGTGCTTCCAGAGATGAAGACCACTGTCTCTCGGAGAC 4012
 Db 243 GTACCTAAGCTATGAGAGTCTGTGCTTCCGAGATGAGAGACCACTGTCTCGGAGAC 302
 QY 4013 TTGAGTCCCGGCTTACCGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 4072
 Db 303 TGAAGGTAACAAGAAAG 362
 QY 4073 TCATCAACAACAAGATGTTCTGCTGCTTCAATCCGACGCTTGAAGTCCAGCTGACT 4132
 Db 363 TCATCAACAACAAGATGTTCTGCTGCTTCAATCCGACGCTTGAAGTCCAGCTGACT 422
 QY 4133 TCTTCATGCGGAGACCGTGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 4192
 Db 423 TCTTCATGCGGAGACCGTGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 481
 QY 4193 TGAAGTACGCACTGATGATGCTGTAAGACAGCTGTGCGGACCTCATTTGAACAAGCTGG 4252
 Db 482 TGAATATGCACTGATGATGCTGTAAGACAGCTGTGCGGACCTCATTTGAACAAGCTGG 541
 QY 4253 AGAGCAAGAACCAACCTTAAGCTGTCTCAGAGAGAGAGAGAGAGAGAGAGAGAGAG 4312
 Db 542 AGAGCAAGAACCAACCTTAAGCTGTCTCAGAGAGAGAGAGAGAGAGAGAGAGAGAG 601
 QY 4313 TGAACCAATGATTTACTTCTCTCTCTCAAGTTCCTCAAGAGAGTGTGCGGAGAGAG 4372
 Db 602 TGAACCAATGATTTACTTCTCTCTCTCAAGTTCCTCAAGAGAGTGTGCGGAGAGAG 661
 QY 4373 TCTTCTCTCTGTTGCTGCTGATTAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 4432
 Db 662 TCTTCTCTCTGTTGCTGCTGATTAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 721
 QY 4433 CGGGAGAGAGCGGCTCTCTCTCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 4492
 Db 722 CGGGAGAGAGCGGCTCTCTCTCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 780
 QY 4493 AAAACCTGTGCTGAGCTGTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 4552
 Db 781 AGACCTGTGCTGAGCTGTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 840
 QY 4553 AGATCCCAACTGATGACACCATCACTCAGTCAAGAGAGAGAGAGAGAGAGAGAGAGAG 4612
 Db 841 AGGTGTTAACTGTGACATCAACAGGTCAAGAGAGAGAGAGAGAGAGAGAGAGAGAG 900
 QY 4613 AGAATGTCTTCTCTCCACCGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 4672
 Db 901 AGAATGTCTTCTCTCCACCGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 960
 QY 4673 GAGTGGGAGCAAGATGATTTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 4717

Db 961 CCGATCCGCGGCTGCT--GCARATGAGACATCACCA 1003

RESULT 12
CA316867 745 bp mRNA linear EST 09-JUL-2003
LOCUS UI-M-FW0-chk-j-22-0-UI.r1 NIH_BMAP_FW0 Mus musculus cDNA clone
DEFINITION IMAGE:6810839.5', mRNA sequence.
ACCESSION CA316867
VERSION CA316867.1 GI:24534991
KEYWORDS EST.
SOURCE Mus musculus (house mouse)
ORGANISM Mus musculus; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE 1 (bases 1 to 745)
AUTHORS NIH-MGC http://mgc.nci.nih.gov/.
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgabbs-remail.nih.gov
Tissue Procurement: Dr. Jim Lin, University of Iowa
cDNA Library Preparation: Dr. M. Bento Soares, University of Iowa
DNA Library Arrayed by: Dr. M. Bento Soares, University of Iowa
DNA Sequencing by: Dr. M. Bento Soares, University of Iowa
Clone Distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LMNL at:
http://image.llnl.gov
This clone was contributed by the Brain Molecular Anatomy Project (BMAP)
Seq primer: pyx-5.
FEATURES
source Location/Qualifiers
1..745
/organism="Mus musculus"
/mol_type="mRNA"
/strain="C57BL/6"
/db_xref="taxon:10090"
/clone="IMAGE:6810839"
/tissue_type="whole brain"
/dev_stage="embryo 13.5, 14.5, 16.5, 17.5 dpc"
/lab_pos="DH10B (T1 phase resistant)"
/clone_1lb="NIH_BMAP_FW0"
/note="Organ: Brain; Vector: pyx-Asc, Site_1: Ecor I; Site_2: Not I; The library was constructed according to Bonaldo, Lennon and Soares, Genome Research, 6:791-806, 1996. Denatured RNA was size fractionated on a 1% agarose gel. First strand cDNA synthesis was primed with oligo-dT primer containing a Not I site. Double strand cDNA was size selected according to mRNA size fraction, ligated with Ecor I adaptor, digested with Not I and then cloned directionally into pyx-Asc vector. The library tag sequence located between the Not I site and the polyA tail is AGCGAGACAG. This library was created for the University Iowa Brain Anatomy Project (BMAP): 'Gene Discovery in the Developing Mouse Nervous System', supported by National Institute of Mental Health (NIMH), Hemlin Chitt, Ph.D., program coordinator."

ORIGIN
Query Match 11.2%; Score 638.4; DB 14; Length 745;
Best Local Similarity 91.0%; Pred. No. 8.7e-124;
Matches 678; Conservative 0; Mismatches 67; Indels 0; Gaps 0;

Db 3867 CAAGAGAGCTTCCGAGCTGACAGCGACATCCATGAGCTGACAGACTTGATG 3926
1 CAAGAGAGCTTCCGAGCTGACAGCAATATTCATGAGCTTACAGGAGCTGATG 60

QY 3927 AGCGGAGATTCGCTCTGAGCTATAGAACTTACACCATGCGGAGCTTTCCAGGAT 3986
Db 61 AGCTGGATTCCTTTCTGAGCTACAGAACTATACAGCGGAGCTTTCCAGGAT 120

QY 3987 TGAAGACCACTGCTCTCCGAGACTTGAAGTCCCGGCTACCGGAGAGCTGTGGA 4046

Db 121 TGAAGACCACTGCTCTCCGAGACTTGAAGTCCCGGCTACCGGAGAGCTTGA 180

QY 4047 GAAAGCTGAAAGCTCTTCCGAGCTCATCAACAAGAGTCTCTGCTTCAT 4106
Db 181 GAAAGCTGAAAGCTCTTCCGAGCTCATCAACAAGAGTCTCTGCTTCAT 240

QY 4107 CCGGAGCTTGAAGTCCCGAGCTGAGTCTTCCATGCGGAGCGGAGAGCTGCT 4166
Db 241 CCGGAGCTTGAAGTCTGAGGAGCTTTCATGCGGAGCGGAGAGCTGCTGCT 300

QY 4167 CATCATGACCTGCTCTGAGAGCAAGCTGAGTACGCACTGATGCTGAGAGCTGCT 4226
Db 301 CATCATGAGCTGCTCTGAGAGCAAGCTGAGTACGCACTGATGCTGAGAGCTGCT 360

QY 4227 GCGGAGCTTGAAGTCCCGAGCTGAGTCTTCCATGCGGAGCGGAGAGCTGCT 4286
Db 361 GCGGAGCTTGAAGTCCCGAGCTGAGTCTTCCATGCGGAGCGGAGAGCTGCT 420

QY 4287 GACTGAGTCAAGTGGCTGAGAGAGTGTGACCAATTGTTACTTCTCTTACAGTT 4346
Db 421 GAGGAGTGGTGGCCGAGAGAGTGTGACTAATTGTTCACTTCTCTTACAGTT 480

QY 4347 CTTGAGAGTGTGCTGGAGAGCTCTTCTCTCTGTTCTGTCATGAGAGAGT 4406
Db 481 CTTGAGAGTGTGCTGGAGAGCTCTTCTCTCTGTTCTGTCATGAGAGAGT 540

QY 4407 GGAAGAGGCGCCATTTGAGAGCTGAGAGAGTGTGACCAATTGTTACTTCTCTTACAGTT 4466
Db 541 GGAAGAGGCGCCATTTGAGAGCTGAGAGAGTGTGACCAATTGTTACTTCTCTTACAGTT 600

QY 4467 GCTCATCCGAGAGCTGAGAGAGTGTGACCAATTGTTACTTCTCTTACAGTT 4526
Db 601 ACTCATCCGAGAGCTGAGAGAGTGTGACCAATTGTTACTTCTCTTACAGTT 660

QY 4527 TGCCAGAGGCGCCATTTGAGAGCTGAGAGAGTGTGACCAATTGTTACTTCTCTTACAGTT 4586
Db 661 TGTTAGAGGCGCCATTTGAGAGCTGAGAGAGTGTGACCAATTGTTACTTCTCTTACAGTT 720

QY 4587 GGAAGAGTGTGAGAGCTGAGAGAGTGTGACCAATTGTTACTTCTCTTACAGTT 4611
Db 721 AGAGAGATCTTGAAGCTGAGAGAGTGTGACCAATTGTTACTTCTCTTACAGTT 745

RESULT 13
BX470595 618 bp mRNA linear EST 04-SEP-2003
LOCUS DFFZp68607110.r1 686 (synonym: hloc3) Homo sapiens cDNA clone
DEFINITION DFFZp68607110.5', mRNA sequence.
ACCESSION BX470595
VERSION BX470595.1 GI:31664922
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE 1 (bases 1 to 618)
AUTHORS Koehler, K., Beyer, A., Mewes, H.W., Weil, B., Amlid, C., Osanger, A., Fodor, G., Han, W., and Wiemann, S.
EST (Koehler, K., Beyer, A., Mewes, H.W., Weil, B., Amlid, C., et al.)
TITLE Unpublished (2003)
JOURNAL Contact: MIPS
COMMENT MIPS
Ingolstaedter Landstr. 1, D-85764 Neuberg, Germany
This is the 5' sequence of the clone insert.
Clone from S. Wiemann, Molecular Genome Analysis, German Cancer Research Center (DKFZ), Email: S.Wiemann@dkfz-heidelberg.de;
Sequenced by BMFZ (Biomedical Research Center at the Heinrich-Heine-University, Dueseldorf/Germany) within the cDNA sequencing consortium of the German Genome Project. No sequence available.
This clone (DKFZp68607110) is available at the RZPD in Berlin.
Please contact the RZPD: Ressourcenzentrum, Heubergweg 6, 14059 Berlin-Charlottenburg, GERMANY; Email: clone@rzpd.de.
Location/Qualifiers

FEATURES


```

source
1. 618
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="DKFZp686907110"
/dev_stage="adult"
/lab_host="DH10B"
/clone_id="686 (synonym: hicc3)"
/note="Vector: pTriplEx2; Site_1: SfilA; Site_2: SfilB;
CDNA-collection"

```

ORIGIN

```

Query Match      10.8%; Score 613; DB 13; Length 618;
Best Local Similarity 100.0%; Pred. No. 1.8e-118;
Matches 613; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

QY 2683 AAGTTTCTGCGCTGAGTGCAGCCCTTTAGTGTGATGATCCCTGAGAAAGATC 2742
DB 6 AAGGTTGCTGCGCTGAGTGCAGCCCTTTAGTGTGATGATCCCTGAGAAAGATC 65
QY 2743 GTGTGTGAGATGGGGAGGCGCAAGCCAGCCAGCATGAGAGCTTGTGAGATCTGCGT 2802
DB 66 GTGTGTGAGATGGGGAGGCGCAAGCCAGCCAGCATGAGAGCTTGTGAGATCTGCGT 125
QY 2803 GCTGTGTGTGCGCTGGAATTCATGCGCGGTCTCTCAAGCTCTATTACTTATGACTG 2862
DB 126 GCTGTGTGTGCGCTGGAATTCATGCGCGGTCTCTCAAGCTCTATTACTTATGACTG 185
QY 2863 ACTCTCTGAGATCTGTAAGCCAGCCGCGGGGCCCATGTCGCGAGGAGCCCAAGTACATC 2922
DB 186 ACTCTCTGAGATCTGTAAGCCAGCCGCGGGGCCCATGTCGCGAGGAGCCCAAGTACATC 245
QY 2923 ACAAGGACCAACCTTGATGCGCGGAAAGCAAGTGTGTGTGATGTTGAAAGAGCCCTGT 2982
DB 246 ACAAGGACCAACCTTGATGCGCGGAAAGCAAGTGTGTGTGATGTTGAAAGAGCCCTGT 305
QY 2983 CTCTTCCAGAGGAGATCTGATCTTCAATGATGTCGTAACACACATCCCTGATAGAGTG 3042
DB 306 CTCTTCCAGAGGAGATCTGATCTTCAATGATGTCGTAACACACATCCCTGATAGAGTG 365
QY 3043 CTAGAGATGAAGGTGTGTGTGAGAGTGAAGCGGCGCAAGATCCACAGAGACTGTCTTT 3102
DB 366 CTAGAGATGAAGGTGTGTGTGAGAGTGAAGCGGCGCAAGATCCACAGAGACTGTCTTT 425
QY 3103 CAGTATGTGAAGACCCACCATGTCGCGGATTTGAGCCAGATGAGAGCTTGTCACTGGA 3162
DB 426 CAGTATGTGAAGACCCACCATGTCGCGGATTTGAGCCAGATGAGAGCTTGTCACTGGA 485
QY 3163 AACACACCATGCGCGTATGCGGGAGCCACCTGAGCTTCAACAGAACCCCAAGTCCGT 3222
DB 486 AACACACCATGCGCGTATGCGGGAGCCACCTGAGCTTCAACAGAACCCCAAGTCCGT 545
QY 3223 GCCAAGCATGAGAGGAGAGAGCATCATCATATCTGAGAGTCTGAAAGCTCTGAGAGT 3282
DB 546 GCCAAGCATGAGAGGAGAGAGCATCATCATATCTGAGAGTCTGAAAGCTCTGAGAGT 605
QY 3283 ACCTGTCAAGCGC 3295
DB 606 ACCTGTCAAGCGC 618

```

```

RESULT 14
BM544169      1027 bp      mRNA      linear      EST 20-FEB-2002
LOCUS        AGENCOURT_6490555 NIH_MGC_125 Homo sapiens CDNA clone IMAGE:5587657
DEFINITION   5', mRNA sequence.
ACCESSION    BM544169
VERSION      BM544169.1 GI:18775199
KEYWORDS     EST.
SOURCE       Homo sapiens (human)
ORGANISM     Eukaryote; Metazoa; Chordata; Cranialia; Vertebrata; Euteleostomi;
              Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

```

```

REFERENCE 1 (bases 1 to 1027)
AUTHORS NIH-MGC http://mgs.nci.nih.gov/.
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
          Email: cgabbs-remail.nih.gov
          Tissue Procurement: Invitrogen
          CDNA Library Preparation: Life Technologies, Inc.
          CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)
          DNA Sequencing by: Agencourt Bioscience Corporation
          Clone distribution: MGC clone distribution information can be
          found through the I.M.A.G.E. Consortium/LNL at:
          http://image.lnl.gov
          Plate: LHAM12357 row: 1 column: 02
          High quality sequence stop: 676.
          Location/Qualifiers

```

FEATURES

source

```

1. 1027
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:5587657"
/lab_host="DH10B"
/clone_id="NIH_MGC_125"
/note="Organ: ovary (pool of 3); Vector: pCMV-SPORT;
Site 1: EcoRV (destroyed); Site 2: NotI; RNA source pool
of three ovaries, from females ranging in age from 38 to
49 yo. Library is oligo-dT primed and directionally cloned
(EcoRV site is destroyed upon cloning). Average insert
size 2.1 kb, insert size range 1-3.5 kb. Library is
normalized and enriched for full-length clones and was
constructed by C. Gruber (Invitrogen). Research Genetics
tracking code 036."

```

ORIGIN

```

Query Match      10.5%; Score 599.2; DB 12; Length 1027;
Best Local Similarity 78.9%; Pred. No. 2e-115;
Matches 764; Conservative 0; Mismatches 198; Indels 6; Gaps 4;

```

```

QY 4280 TCAGAGAGACTGAGTCAAGTGTGAGAGATGTCACCAATGTTACTTTCTCTCT 4339
DB 12 TCCTGGATGTGATGATGATGAGAGATGTCATCACTGTTCACTTCTCTCTGT 71
QY 4340 ACAAGTTCTCAAGAGTGTCTGCGGAGCCCTCTTCTCCCTGTTCTGTCATCAAGC 4399
DB 72 ATAAATTCCTCAAGAGTGTGCGTGGAGCGCTGTTCAATGCTACTGTGCGCATCAAGC 131
QY 4400 AGCAGATGAGAGGGGCCCATTTGACGACATCAAGGCGGAGCCGCTACTCTTGAAGC 4459
DB 132 AGCAGATGAGAGGGGCCCATTTGACGACATCAAGGCGGAGCGCTACTCTTGAAGT 191
QY 4460 AGCAGATGATCCGCGCAGACATTTGATCAAAAACCTGTGCTGAGTGTGTCAAGC 4519
DB 192 AGCAGATGATCCGCGCAGACATTTGATCAAAAACCTGTGCTGAGTGTGTCAAGC 251
QY 4520 CAGCATGCGCAAGCGCCGAGAGTCCAGTAAAGATCTCACTGTGACACCATCACTC 4579
DB 252 CTGGAATGAGATTCACCTGAGAGTCCGCTGAAGGGGCTGCACTGACAGCGTCAACC 311
QY 4580 AGTCAAGAGAGATTTCTGATGACATCTTCAAGATGTGCTTGTCTCCACCGGCCA 4639
DB 312 AGGCAAGAGAGATCTGATGACATCTTCAAGATGTGCTTGTCTCCACCGGCCA 371
QY 4640 AAGCTGAGATTTGATGATGATGAGTGGGCAAGAGATGAGGAGATGATCTTGAGG 4699
DB 372 AGGCGCGGACATGAGATCTGAGTGGGCGGAGGCGGAGATGATCTTGAGG 431
QY 4700 ATGAGACATCACCAACCAAGATTTGAGATGTTGAAGCACTGACATCTGGCCACT 4759
DB 432 ACAGAGAGCTCACCAACCAAGATTTGAGATGTTGAAGCACTGACATCTGGCCACT 491
QY 4760 ACCAGGTGACAGATGTTCTCGTGTGATTTAATGTCCAGAGGTGACAGCGCTATAGC 4819
DB 492 ACCAGGTGACAGAGGTTCTCGTGTGATTTAATGTCCAGAGGTGACAGCGCTATAGCA 551

```

QY 4820 CAGTGAACACTCCACCGTCTCCAGAGACTCAGCAAGTAATGAAAACATGATCCGGT 4879
DB 552 TCTCCAACTCTCCACCTTCAACCAAG---TCCCTCAGCAGATACGAGCAATGCTGCCA 608
QY 4880 AACAGGAG 4939
DB 609 CGGCGAG 668
QY 4940 GTGAGTCAAGATGTGGACCTTGTGAAGAACAGAGAGAGAGAGAGAGAGAGAGAG 4999
DB 669 GCGGAG 727
QY 5000 ACCGAG 5059
DB 728 ACCGAG 787
QY 5060 CACTGAG 5118
DB 788 CACTGAG 847
QY 5119 TCTGAG 5177
DB 848 TCGAG 907
QY 5178 ACATGAG 5237
DB 908 GACAG 967
QY 5238 GTTTTGGG 5245
DB 968 CTTTCTGG 975

RESULT 15
BU149880 882 bp mRNA linear EST 03-SRP-2002
LOCUS BU149880
DEFINITION AGENCOURT 8095739 NIH_MGC_112 Homo sapiens cDNA clone IMAGE:6088019
5' mRNA sequence.
ACCESSION BU149880
VERSION BU149880.1 GI:22663412
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Cranialia; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
REFERENCE 1 (bases 1 to 882)
NIH-MGC http://mgs.nci.nih.gov/.
National Institutes of Health, Mammalian Gene Collection (MGC)
Unpublished (1999)
Contact: Robert Strausberg, Ph.D.
Email: gcraps-remail.nih.gov
Tissue Procurement: DCTD/DTF
CDNA Library Preparation: Rubin Laboratory
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LINL)
DNA Sequencing by: Agencourt Bioscience Corporation
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LINL at:
http://image.lnl.gov
Plate: LINC2325 row: 1 column: 12
High quality sequence stop: 631.
Location/Qualifiers
1.882
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:6088019"
/tissue_type="melanotic melanoma, cell line"
/lab_host="DH10B (phage-resistant)"
/clone_id="NIH MGC 112"
/note="Organ: skin; Vector: pOTB7; Site 1: XhoI; Site 2:
EcoRI; cDNA made by oligo-dT priming. Directional cloning
strategy."</p></div>
<div data-bbox="742 522 816 943" data-label="Text">
<p>GGCAGAG(G). Library constructed by Ling Hong in the

Laboratory of Gerald M. Rubin (University of California,

Berkeley) using ZAP-cDNA synthesis kit (Stratagene) and

Superscript II RT (Life Technologies). Note: this is a

NH_MGC library."</p>
</div>
<div data-bbox="30 522 735 946" data-label="Text">
<p>ORIGIN

Query March 10.2%; Score 578; DB 13; Length 882;

Best Local Similarity 78.5%; Pred. No. 5,6e-111;

Matches 692; Conservative 0; Mismatches 190; Indels 0; Gaps 0;

4363 GGGAG 4422

1 GGGAG 60

4423 GAGGAG 4482

61 GATGAG 120

4483 ATTGAG 4542

121 ATGAG 180

4543 GTCCAG 4602

181 ATCCAG 240

4603 GCCATCTTCAAG 4662

241 GCGGTATTAAG 300

4663 TGCGAG 4722

301 TGCGAG 360

4723 GAG 4782

361 GAGGAG 420

4783 GTGAG 4842

421 GTGAG 480

4843 AG 4902

481 CGAG 540

4903 CGTCAAG 4962

541 CGGTCCGAG 600

4963 GTGAAG 5022

601 GTGAAG 660

5023 GAAATCTACCTGAG 5082

661 GAGATCTACCTGAG 720

5083 CTCCTTGAAG 5142

721 TTGTTTGAAG 780

5143 TACATGTTTGAAG 5202

781 TACATGTTTGAAG 840

5203 CGCATCTTGAAG 5244

841 CGGACAG 882</p>
</div>
<div data-bbox="0 522 19 766" data-label="Page-Footer">
<p>Search completed: February 19, 2004, 23:13:56</p>
</div>

Mon Feb 23 09:34:03 2004

us-09-964-956-12.rst

Job time : 8756 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Comugen Ltd.

score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

OM nucleic - nucleic search, using sw model

Run on: February 19, 2004, 23:18:35 / Search time 1415 Seconds

(without alignments)
17475.395 Million cell updates/sec

Title: US-09-964-956-12

Perfect score: 5691 1 atgaaagccatgccttgaaa.....gcttagacagctgaataaa 5691

Sequence: 1 atgaaagccatgccttgaaa.....gcttagacagctgaataaa 5691

Scoring table: OLIGO NUC
Gapop 60.0 , Gapext 60.0

Searched: 3470272 seqs, 21671516995 residues

Word size : 0
Total number of hits satisfying chosen parameters: 6940544

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Listing first 1000 summaries

Database :

GenBank: 1: gb_ba: 2: gb_htg: 3: gb_in: 4: gb_cm: 5: gb_ov: 6: gb_pat: 7: gb_ph: 8: gb_pl: 9: gb_pr: 10: gb_ro: 11: gb_sbs: 12: gb_sy: 13: gb_un: 14: gb_vl: 15: em_ba: 16: em_fun: 17: em_hum: 18: em_in: 19: em_mu: 20: em_om: 21: em_or: 22: em_ov: 23: em_pat: 24: em_ph: 25: em_pl: 26: em_ro: 27: em_sbs: 28: em_un: 29: em_vl: 30: em_htg_hum: 31: em_htg_in: 32: em_htg_juv: 33: em_htg_mus: 34: em_htg_pln: 35: em_htg_rnd: 36: em_htg_mam: 37: em_htg_vrt: 38: em_sy: 39: em_htgo_hum: 40: em_htgo_mus: 41: em_htgo_other:

Pred. No. is the number of results predicted by chance to have a

Result No.	Score	Query Match	Length	DB ID	Description
1	5691	100.0	5691	6	AX503782 Sequence
2	5532	97.2	6367	6	AX492982 Sequence
3	2237	39.3	5986	9	AB046770 Homo sapi
4	1371	24.1	2015	9	BC028744 Homo sapi
5	1371	24.1	3310	9	AK123428 Homo sapi
6	1188	20.9	179145	9	AC009785 Homo sapi
7	1108	19.5	3666	6	AX833678 Sequence
8	1108	19.5	3666	9	AK095606 Homo sapi
9	1106	19.4	188090	9	AC011625 Homo sapi
10	1068	18.8	179592	2	AC009723 Homo sapi
11	878	15.4	3556	6	HSMB805617 Homo sapi
12	793	13.9	2597	6	AX574564 Sequence
13	793	13.9	2597	9	AY358850 Homo sapi
14	706	12.4	1029	9	HSMB802037 Homo sapi
15	509	8.9	74655	2	AC025595 Homo sapi
16	491	8.6	601	6	AX088116 Sequence
17	479	8.4	3560	6	AX747045 Homo sapi
18	479	8.4	3560	9	AK091538 Homo sapi
19	272	4.8	198470	9	AC018643 Homo sapi
20	241	4.2	171557	2	AC147275 Homo sapi
21	216	3.8	185272	9	AC105443 Homo sapi
22	134	2.4	151153	2	AC009364 Homo sapi
23	123	2.4	183538	2	AC147272 Homo sapi
24	123	2.2	171557	2	AC147276 Homo sapi
25	74	1.3	74655	2	AC025595 Homo sapi
26	61	1.1	247656	2	AC112331 Rattus no
27	61	1.1	274459	2	AC093959 Rattus no
28	61	1.1	284400	2	AC096840 Rattus no
29	53	0.9	7022	10	AB073228 Mus muscu
30	53	0.9	172060	2	AC102171 Mus muscu
31	49	0.9	6039	9	HSSEKXENB Homo sapi
32	48	0.8	14560	2	BX664737 Homo sapi
33	48	0.8	160656	2	AL592072 Homo sapi
34	48	0.8	219447	9	HUMFLNG6PD Homo sapi
35	47	0.8	4445	9	AK126101 Homo sapi
36	47	0.8	5263	9	HSNOVPROT Homo sapien
37	47	0.8	5895	6	AX686475 Sequence
38	47	0.8	6147	6	AX704766 Sequence
39	47	0.8	6693	6	BD183298 Novel gen
40	47	0.8	91240	9	AC130566 Homo sapi
41	47	0.8	152135	9	AC011199 Homo sapi
42	47	0.8	221370	2	AC126205 Rattus no
43	47	0.8	22898	2	AC131584 Rattus no
44	45	0.8	465	11	HSCE7605 R. sapiens
45	44	0.8	172060	2	AC102171 Mus muscu
46	44	0.8	238662	2	AC102156 Mus muscu
47	39	0.7	1817	9	BC006193 Homo sapi
48	39	0.7	2875	9	BC009343 Homo sapi
49	39	0.7	3985	9	HSOCHPROT Homo sapien
50	39	0.7	6263	9	AB007932 Homo sapi
51	39	0.7	6387	6	AX405895 Sequence
52	39	0.7	118429	9	AL158017 Homo sapi
53	39	0.7	198906	9	AL356275 Human DNA
54	38	0.7	776	4	AF176418 Bos tauru
55	38	0.7	6466	10	D86948 Mus musculu
56	38	0.7	211217	2	AC123387 Rattus no
57	35	0.6	218548	2	AC132672 Rattus no
58	35	0.6	228751	2	AX644286 Dario rer
59	35	0.6	229222	2	AX102141 Mus muscu
60	35	0.6	230199	5	EX537358 Zebrafish
61	35	0.6	251269	2	AC145193 Gallus ga
62	35	0.6	260656	2	EX640582 Dario rer
63	34	0.6	5982	10	D86950 Mus musculu
64	34	0.6	188873	2	AC073658 Mus muscu
65	34	0.6	198653	2	AC091474 Mus muscu

66	34	0.6	200322	2	AL808145	AL808145 Mus muscu	139	21	0.4	2815	10	BC011069
67	34	0.6	203189	10	AL807376	AL807376 Mouse DNA	140	21	0.4	3078	10	WMPCTA1A
68	34	0.6	214505	2	AC025784	AC025784 Mus muscu	141	21	0.4	3111	10	RNU36444
69	34	0.6	220228	2	AC094668	AC094668 Rattus no	142	21	0.4	5249	10	AC122467
70	33	0.6	146405	1	AC122795	AC122795 Mus muscu	143	21	0.4	46794	2	AC114683
71	33	0.6	255914	2	AC105876	AC105876 Rattus no	144	21	0.4	60111	2	AC117617
72	32	0.6	1755	10	BC043322	BC043322 Mus muscu	145	21	0.4	65513	2	AC103748
73	32	0.6	4320	10	AK128133	AK128133 Mus muscu	146	21	0.4	66118	9	AL162259
74	32	0.6	8466	10	AB072381	AB072381 Mus muscu	147	21	0.4	73947	9	AL445248
75	32	0.6	109148	2	AC138439	AC138439 Takifugu	148	21	0.4	83559	9	AC044786
76	31	0.5	208593	2	AC140383	AC140383 Mus muscu	149	21	0.4	98861	9	AL136961
77	31	0.5	229392	2	AC112372	AC112372 Rattus no	150	21	0.4	99512	9	AC004606
78	31	0.5	238662	2	AC102156	AC102156 Mus muscu	151	21	0.4	101458	2	AP000680
79	31	0.5	255561	2	AC121624	AC121624 Rattus no	152	21	0.4	109602	2	AC131780
80	29	0.5	582	5	AB055678	AB055678 Dantio rer	153	21	0.4	109640	10	AC135241
81	29	0.5	768	6	AB055678	AB055678 Dantio rer	154	21	0.4	109624	10	AC144935
82	29	0.5	768	6	BD145964	BD145964 Primer fo	155	21	0.4	110000	2	AC109940
83	29	0.5	1012	6	BC032125	BC032125 Homo sapi	156	21	0.4	115145	10	AL807240
84	29	0.5	1767	6	AB078694	AB078694 Sequence	157	21	0.4	129414	2	AC143614
85	29	0.5	1767	6	BD157390	BD157390 Primer fo	158	21	0.4	134506	9	HS357116
86	29	0.5	2294	9	AK094367	AK094367 Homo sapi	159	21	0.4	135638	1	AF484556
87	29	0.5	2294	9	AK094367	AK094367 Homo sapi	160	21	0.4	137908	10	AL607024
88	29	0.5	4196	9	AK128612	AK128612 Homo sapi	161	21	0.4	142696	2	AC143051
89	26	0.5	26	6	AX503877	AX503877 Sequence	162	21	0.4	147138	10	AL672094
90	26	0.5	6163	5	XELPLEX	D38175 Xenopus lae	163	21	0.4	151174	2	AC016189
91	26	0.5	75429	2	AC101416	AC101416 Mus muscu	164	21	0.4	154896	2	AC110481
92	25	0.4	6330	6	BC056475	BC056475 Mus muscu	165	21	0.4	158170	8	AC093180
93	25	0.4	6330	10	AB094949	AB094949 Mus muscu	166	21	0.4	160989	8	OSJN00089
94	25	0.4	6330	10	AB094949	AB094949 Mus muscu	167	21	0.4	165776	2	AC109157
95	25	0.4	6893	10	AK122289	AK122289 Mus muscu	168	21	0.4	166894	8	OSJN00077
96	25	0.4	212744	2	AC106326	AC106326 Rattus no	169	21	0.4	168639	2	AC121229
97	25	0.4	216192	10	AC111280	AC111280 Rattus no	170	21	0.4	171368	9	HS095738
98	25	0.4	228202	5	BX284664	BX284664 Zebrafish	171	21	0.4	172029	9	AC104416
99	24	0.4	1486	6	OCHEMOPEX	X16429 Oryzolaes	172	21	0.4	172506	2	AC105743
100	23	0.4	190	6	AX904782	AX904782 Sequence	173	21	0.4	175745	2	BX530060
101	23	0.4	1758	9	BD040315	BD040315 Sequence	174	21	0.4	175849	2	AL512366
102	23	0.4	147686	10	HSW802538	AL162013 Homo sapi	175	21	0.4	175955	2	AC131778
103	23	0.4	157250	2	AC027264	AC027264 Homo sapi	176	21	0.4	176329	2	AC118047
104	23	0.4	164057	9	AL590138	AL590138 Human DNA	177	21	0.4	178525	9	AC018612
105	23	0.4	169305	9	AC021036	AC021036 Homo sapi	178	21	0.4	189258	2	AC024021
106	23	0.4	183353	10	AL935878	AL935878 Mouse DNA	180	21	0.4	183133	10	AC101851
107	23	0.4	190748	9	AC010969	AC010969 Homo sapi	181	21	0.4	185973	2	AC128212
108	23	0.4	217392	2	AC107414	AC107414 Rattus no	182	21	0.4	189258	2	AC122985
109	23	0.4	218448	2	BX324430	BX324430 Dantio rer	183	21	0.4	190104	2	AC147396
110	23	0.4	229132	2	AC112047	AC112047 Rattus no	184	21	0.4	192729	2	AC124709
111	23	0.4	253343	10	AC102566	AX503876 Mus muscu	185	21	0.4	195611	9	AC073957
112	22	0.4	6178	22	AX503876	AX503876 Sequence	186	21	0.4	196204	9	AC073957
113	22	0.4	143406	10	AC092857	AC092857 Rattus no	187	21	0.4	198197	10	AC125035
114	22	0.4	174928	10	AC121875	AC121875 Mus muscu	188	21	0.4	198197	10	AC125035
115	22	0.4	182432	9	AL161771	AL161771 Human DNA	189	21	0.4	204426	2	AC108778
116	22	0.4	190135	2	AC142039	AC142039 Mus muscu	190	21	0.4	205968	9	AC073912
117	22	0.4	197333	2	AC120001	AC120001 Mus muscu	191	21	0.4	223465	2	AC131813
118	22	0.4	202391	10	AC110040	AC110040 Mus muscu	192	21	0.4	223465	2	AC131813
119	22	0.4	205587	2	AC139761	AC139761 Mus muscu	193	21	0.4	223465	2	AC131813
120	22	0.4	207228	2	AC109215	AC109215 Mus muscu	194	21	0.4	223465	2	AC131813
121	22	0.4	228678	2	AC110740	AC110740 Mus muscu	195	21	0.4	223465	2	AC131813
122	22	0.4	230329	10	AC142070	AC142070 Rattus no	196	21	0.4	223465	2	AC131813
123	22	0.4	232691	2	AC127871	AC127871 Mus muscu	197	21	0.4	223465	2	AC131813
124	22	0.4	234882	2	AC127871	AC127871 Mus muscu	198	21	0.4	223465	2	AC131813
125	22	0.4	237777	2	AC131632	AC131632 Rattus no	199	21	0.4	223465	2	AC131813
126	22	0.4	244795	2	AC105641	AC105641 Rattus no	200	21	0.4	223465	2	AC131813
127	22	0.4	246196	2	AC105641	AC105641 Rattus no	201	21	0.4	223465	2	AC131813
128	22	0.4	254380	2	AC129440	AC129440 Rattus no	202	21	0.4	223465	2	AC131813
129	22	0.4	273235	1	AX572606	AX572606 Rhodosphe	203	21	0.4	223465	2	AC131813
130	22	0.4	349841	1	AX503878	AX503878 Sequence	204	21	0.4	223465	2	AC131813
131	22	0.4	21	6	AF540386	AF540386 Mus muscu	205	21	0.4	223465	2	AC131813
132	21	0.4	1874	10	RATGABAAT	RATGABAAT Rattus norv	206	21	0.4	223465	2	AC131813
133	21	0.4	2134	10	BC013663	BC013663 Mus muscu	207	21	0.4	223465	2	AC131813
134	21	0.4	2591	10	RNGABAAS	RNGABAAS Rattus norv	208	21	0.4	223465	2	AC131813
135	21	0.4	2689	10	BC062112	BC062112 Mus muscu	209	21	0.4	223465	2	AC131813
136	21	0.4	2689	10	BC062112	BC062112 Mus muscu	210	21	0.4	223465	2	AC131813
137	21	0.4	2689	10	BC062112	BC062112 Mus muscu	211	21	0.4	223465	2	AC131813
138	21	0.4	2689	10	BC062112	BC062112 Mus muscu	211	21	0.4	223465	2	AC131813

C 212	21	0.4	244343	2	AC106516	Rattus no	C 285	20	0.4	2660	9	AK022651	AK022651 Homo sapi
C 213	21	0.4	245892	2	AC096884	Rattus no	C 286	20	0.4	2166	9	BC004542	BC004542 Homo sapi
C 214	21	0.4	247645	2	AC118405	Rattus no	C 287	20	0.4	2291	9	BC011892	BC011892 Homo sapi
C 215	21	0.4	250415	2	AC097418	Rattus no	C 288	20	0.4	2389	9	AF464935	AF464935 Homo sapi
C 216	21	0.4	250627	2	AC097031	Rattus no	C 289	20	0.4	2444	6	E36720	E36720 Novel trans
C 217	21	0.4	254288	2	AC126112	Rattus no	C 290	20	0.4	2522	6	AR339211	AR339211 Sequence
C 218	21	0.4	254708	2	AC111761	Rattus no	C 291	20	0.4	2533	10	BC004762	BC004762 Homo sapi
C 219	21	0.4	254708	2	AC115307	Rattus no	C 292	20	0.4	2564	6	AK1746754	AK1746754 Homo sapi
C 220	21	0.4	262886	2	AC121209	Rattus no	C 293	20	0.4	2564	6	AK091056	AK091056 Homo sapi
C 221	21	0.4	265567	2	AC096635	Rattus no	C 294	20	0.4	2775	10	AB041607	AB041607 Mus muscu
C 222	21	0.4	271040	2	AC121473	Rattus no	C 295	20	0.4	2935	10	BC003293	BC003293 Mus muscu
C 223	21	0.4	276800	1	SC093115	Streptomy	C 296	20	0.4	2971	6	BD127371	BD127371 Primer fo
C 224	21	0.4	279593	2	AC114015	Rattus no	C 297	20	0.4	2971	6	BD127371	BD127371 Primer fo
C 225	21	0.4	280025	2	AC108655	Rattus no	C 298	20	0.4	3038	9	AK126384	AK126384 Homo sapi
C 226	21	0.4	283782	2	AC096538	Rattus no	C 299	20	0.4	3184	10	BC051045	BC051045 Mus muscu
C 227	21	0.4	294396	2	AC105485	Rattus no	C 300	20	0.4	3504	9	AK056543	AK056543 Homo sapi
C 228	21	0.4	316005	2	AC111792	Rattus no	C 301	20	0.4	3726	9	AK025701	AK025701 Homo sapi
C 229	21	0.4	346542	2	AC120727	Rattus no	C 302	20	0.4	3849	6	AX058889	AX058889 Sequence
C 230	21	0.4	347368	2	AC096620	Rattus no	C 303	20	0.4	4125	8	CRE427966	CRE427966 Chlamydom
C 231	20	0.4	207	9	HS62A1R	H. sapiens	C 304	20	0.4	4355	10	BC054788	BC054788 Mus muscu
C 232	20	0.4	242	9	HS62A1R	H. sapiens m	C 305	20	0.4	4378	10	BC053092	BC053092 Mus muscu
C 233	20	0.4	272	9	HS179B11F	Z54966 H. sapiens C	C 306	20	0.4	4725	10	AF190578	AF190578 Homo sapi
C 234	20	0.4	311	6	AR251662	Sequence	C 307	20	0.4	5310	10	BC007481	BC007481 Homo sapi
C 235	20	0.4	318	6	BD004150	BD004150 5'SST and	C 308	20	0.4	6252	6	AX088110	AX088110 Sequence
C 236	20	0.4	324	6	AX408405	AX408405 Sequence	C 309	20	0.4	6252	6	AX210612	AX210612 Sequence
C 237	20	0.4	328	6	AX260802	AX260802 Sequence	C 310	20	0.4	6252	6	AX35107	AX35107 Human mRN
C 238	20	0.4	332	10	AF029886	AF029886 Rattus no	C 311	20	0.4	6252	6	AB002313	AB002313 Streptococ
C 239	20	0.4	381	9	HS430739	HS430739 Homo sapi	C 312	20	0.4	8960	1	AF173226	AF173226 Pseudomon
C 240	20	0.4	382	6	AX895484	AX895484 Sequence	C 313	20	0.4	11158	1	AE004452	AE004452 Streptococ
C 241	20	0.4	382	6	BD031017	BD031017 Sequence	C 314	20	0.4	12744	1	AF204401	AF204401 Xylella I
C 242	20	0.4	412	6	AX895483	AX895483 Sequence	C 315	20	0.4	13714	1	AE003997	AE003997 Xylella I
C 243	20	0.4	412	6	BD031016	BD031016 Sequence	C 316	20	0.4	13842	6	BD217185	BD217185 Mouse DNA
C 244	20	0.4	419	6	AX396560	AX396560 Sequence	C 317	20	0.4	18979	10	BX842567	BX842567 Human DNA
C 245	20	0.4	456	6	AX895482	AX895482 Sequence	C 318	20	0.4	23518	9	AL589735	AL589735 Human DNA
C 246	20	0.4	456	6	BD031015	BD031015 Sequence	C 319	20	0.4	27950	9	BX649592	BX649592 Human DNA
C 247	20	0.4	462	6	BD65166	BD65166 Compounds	C 320	20	0.4	29074	3	LMF05213T	LMF05213T Leishmani
C 248	20	0.4	462	6	AR401152	AR401152 Sequence	C 321	20	0.4	32539	9	HS1LC2	HS1LC2 Human DNA
C 249	20	0.4	462	6	AX192598	AX192598 Sequence	C 322	20	0.4	37948	1	AF079138	AF079138 Streptococ
C 250	20	0.4	486	6	AX895485	AX895485 Sequence	C 323	20	0.4	38506	6	BD217173	BD217173 DNA encod
C 251	20	0.4	486	6	BD031018	BD031018 Sequence	C 324	20	0.4	38506	6	BD232534	BD232534 Recombina
C 252	20	0.4	506	6	AX804284	AX804284 Sequence	C 325	20	0.4	38506	6	AR271680	AR271680 Sequence
C 253	20	0.4	615	6	AX884867	AX884867 Sequence	C 326	20	0.4	38506	6	AR277664	AR277664 Sequence
C 254	20	0.4	615	6	BD024477	BD024477 Sequence	C 327	20	0.4	40548	8	AB011480	AB011480 Arabidops
C 255	20	0.4	616	6	BD229563	BD229563 Human gen	C 328	20	0.4	48662	3	AC024755	AC024755 Genothab
C 256	20	0.4	635	8	AF362041	AF362041 Dendrobii	C 329	20	0.4	52573	2	AC136259	AC136259 Rattus no
C 257	20	0.4	670	6	AX899105	AX899105 Sequence	C 330	20	0.4	59657	9	AC074013	AC074013 Homo sapi
C 258	20	0.4	670	6	BD034638	BD034638 Sequence	C 331	20	0.4	62297	10	AY115107	AY115107 Mus muscu
C 259	20	0.4	689	6	AX899102	AX899102 Sequence	C 332	20	0.4	67173	9	AL359962	AL359962 Human DNA
C 260	20	0.4	689	6	BD034635	BD034635 Sequence	C 333	20	0.4	70282	9	AP001505	AP001505 Homo sapi
C 261	20	0.4	700	6	AX410652	AX410652 Sequence	C 334	20	0.4	74194	9	AL566449	AL566449 Human DNA
C 262	20	0.4	700	6	BC006791	BC006791 Homo sapi	C 335	20	0.4	74595	9	AC130276	AC130276 Homo sapi
C 263	20	0.4	700	9	HS012404	HS012404 Human Csa-1	C 336	20	0.4	77758	8	AB016885	AB016885 Arabidops
C 264	20	0.4	709	9	BC011366	BC011366 Homo sapi	C 337	20	0.4	79340	8	AP003956	AP003956 Oryza sat
C 265	20	0.4	741	6	BD189960	BD189960 Tumor ant	C 338	20	0.4	79884	9	AC109440	AC109440 Homo sapi
C 266	20	0.4	741	6	AB082926	AB082926 Homo sapi	C 339	20	0.4	85440	2	AC034264	AC034264 Mus muscu
C 267	20	0.4	744	6	AX899094	AX899094 Sequence	C 340	20	0.4	88421	6	AX417445	AX417445 Sequence
C 268	20	0.4	744	6	BD034627	BD034627 Sequence	C 341	20	0.4	88588	9	AC010453	AC010453 Homo sapi
C 269	20	0.4	750	6	BD220723	BD220723 Human gen	C 342	20	0.4	89743	10	AC003062	AC003062 Mouse Chr
C 270	20	0.4	824	6	AX867747	AX867747 Sequence	C 343	20	0.4	92580	2	AL591480	AL591480 Human DNA
C 271	20	0.4	824	6	BD147479	BD147479 Primer fo	C 344	20	0.4	92580	2	AC016302	AC016302 Mouse DNA
C 272	20	0.4	1396	6	S76730	S76730 KMI-clone M	C 345	20	0.4	95593	2	BX470092	BX470092 Mouse DNA
C 273	20	0.4	1403	10	BC024509	BC024509 Mus muscu	C 346	20	0.4	98522	10	AC110929	AC110929 (8 of
C 274	20	0.4	1413	9	BT006887	BT006887 Homo sapi	C 347	20	0.4	99108	9	AC008665	AC008665 Homo sapi
C 275	20	0.4	1413	12	BT007594	BT007594 Synthetic	C 348	20	0.4	99108	9	AL928807	AL928807 Mus muscu
C 276	20	0.4	1618	10	MX058887	MX058887 Mus muscu	C 349	20	0.4	100719	9	EX255925	EX255925 Human DNA
C 277	20	0.4	1636	6	AR175265	AR175265 Sequence	C 350	20	0.4	101305	10	AL6633025	AL6633025 Mouse DNA
C 278	20	0.4	1652	10	BC027096	BC027096 Mus muscu	C 351	20	0.4	106323	9	AL513477	AL513477 Human DNA
C 279	20	0.4	1782	10	AF077738	AF077738 Rat notwe	C 352	20	0.4	108015	10	AL645744	AL645744 Mouse DNA
C 280	20	0.4	1860	9	BC061993	BC061993 Rattus no	C 353	20	0.4	108795	9	AC114483	AC114483 Homo sapi
C 281	20	0.4	1876	10	HSMB00181	HSMB00181 Homo sapi	C 354	20	0.4	110000	2	AC013623	AC013623 Mus muscu
C 282	20	0.4	1946	6	AK025415	AK025415 Homo sapi	C 355	20	0.4	110000	2	AC107201	AC107201 (3 of
C 283	20	0.4	2060	6	AX879531	AX879531 Sequence	C 356	20	0.4	110000	2	AC109503	AC109503 (4 of
C 284	20	0.4	2060	6	BD157886	BD157886 Primer fo	C 357	20	0.4	110000	2	AC142490	AC142490 (2 of

[illegible]

C 504	20	0.4	222182	2	AC123412	C 577	20	0.4	275900	2	AC123150	AC123150	Rattus no
C 505	20	0.4	222589	10	AC113364	C 578	20	0.4	281085	2	AC123455	AC123455	Rattus no
C 506	20	0.4	222639	2	AC133083	C 579	20	0.4	284347	2	AC135290	AC135290	Mus muscu
C 507	20	0.4	223911	2	AC136740	C 580	20	0.4	285152	2	AC116112	AC116112	Mus muscu
C 508	20	0.4	224138	10	AC133573	C 581	20	0.4	286451	2	AC115130	AC115130	Rattus no
C 509	20	0.4	224764	2	AC122283	C 582	20	0.4	289103	2	AC126288	AC126288	Rattus no
C 510	20	0.4	224859	2	AC073800	C 583	20	0.4	290753	2	AC123256	AC123256	Rattus no
C 511	20	0.4	225012	2	AC113462	C 584	20	0.4	292100	1	SC0939121	SC0939121	Strepomy
C 512	20	0.4	225403	2	AC140451	C 585	20	0.4	300029	1	AE016842	AE016842	Salmonell
C 513	20	0.4	225509	2	AC073782	C 586	20	0.4	303688	10	AE016754	AE016754	Mus muscu
C 514	20	0.4	225541	2	AC126700	C 587	20	0.4	312626	2	AC109713	AC109713	Rattus no
C 515	20	0.4	225782	2	AC109529	C 588	20	0.4	325265	2	AC118802	AC118802	Rattus no
C 516	20	0.4	227848	2	AC123606	C 589	20	0.4	325265	2	AC118802	AC118802	Rattus no
C 517	20	0.4	228051	2	AC097434	C 590	20	0.4	340000	9	HS21C101	HS21C101	Homo sapi
C 518	20	0.4	228178	2	AC111655	C 591	20	0.4	348764	1	BX569689	BX569689	Synechoco
C 519	20	0.4	228413	2	AC132453	C 592	20	0.3	348764	1	BTL29486	BTL29486	Bos taurus
C 520	20	0.4	228852	2	AC131526	C 593	19	0.3	382	11	AY254001	AY254001	Sus scrofa
C 521	20	0.4	229909	2	AC111310	C 594	19	0.3	426	9	HS4240590	HS4240590	Homo sapi
C 522	20	0.4	230188	2	AC091604	C 595	19	0.3	429	10	F257137S07	F257137S07	Mus muscu
C 523	20	0.4	233180	2	AC134135	C 596	19	0.3	503	5	AB062427	AB062427	Anguilla
C 524	20	0.4	233510	2	AC136569	C 597	19	0.3	503	6	BD012143	BD012143	Novel pep
C 525	20	0.4	233718	2	AC094024	C 598	19	0.3	529	11	G58051	G58051	SHGC-103874
C 526	20	0.4	234053	2	AC128282	C 599	19	0.3	539	11	G97698	G97698	S208P6575FE
C 527	20	0.4	234300	2	AC118431	C 600	19	0.3	545	4	AF381310	AF381310	Oryzathor
C 528	20	0.4	235050	2	AC096077	C 601	19	0.3	553	1	AF448095	AF448095	Synechoco
C 529	20	0.4	235427	2	AC140082	C 602	19	0.3	560	1	AF245157	AF245157	Synechoco
C 530	20	0.4	236159	2	AC130109	C 603	19	0.3	560	1	AF245158	AF245158	Synechoco
C 531	20	0.4	237304	2	AC096114	C 604	19	0.3	561	1	AF448096	AF448096	Synechoco
C 532	20	0.4	238007	10	AC097745	C 605	19	0.3	565	1	AF245125	AF245125	Synechoco
C 533	20	0.4	239628	2	AC128574	C 606	19	0.3	567	1	AF245164	AF245164	Synechoco
C 534	20	0.4	240203	2	AC126562	C 607	19	0.3	570	11	HSC01A04	HSC01A04	H. sapiens
C 535	20	0.4	240362	2	AC132653	C 608	19	0.3	572	1	AF245155	AF245155	Synechoco
C 536	20	0.4	241781	2	AC122087	C 609	19	0.3	588	6	AR093091	AR093091	Sequence
C 537	20	0.4	242000	2	AC127751	C 610	19	0.3	588	6	BD009915	BD009915	BH3 inter
C 538	20	0.4	242482	2	AC115149	C 611	19	0.3	612	1	SSU52342	SSU52342	Synechococ
C 539	20	0.4	242531	10	AC108805	C 612	19	0.3	612	1	SSU52343	SSU52343	Foot-and-
C 540	20	0.4	242662	2	AC080020	C 613	19	0.3	648	14	AF3677124	AF3677124	Synechoco
C 541	20	0.4	243047	2	AC126579	C 614	19	0.3	658	9	AY449016	AY449016	Autos asa
C 542	20	0.4	243369	2	AC117152	C 615	19	0.3	658	9	AY449075	AY449075	Autos asa
C 543	20	0.4	243370	10	AC122324	C 616	19	0.3	658	9	AY449165	AY449165	Ateles fu
C 544	20	0.4	243770	2	AC120763	C 617	19	0.3	739	6	BD125041	BD125041	Primer fo
C 545	20	0.4	244814	2	AC098223	C 618	19	0.3	739	6	BD126697	BD126697	Primer fo
C 546	20	0.4	245790	2	AC130142	C 619	19	0.3	759	11	BV044442	BV044442	S212P6045
C 547	20	0.4	246238	2	AC120998	C 620	19	0.3	779	11	BV063339	BV063339	Sequence
C 548	20	0.4	246921	2	AC107248	C 621	19	0.3	791	6	AR271766	AR271766	Sequence
C 549	20	0.4	247108	2	AC123333	C 622	19	0.3	791	6	MM075506	MM075506	Mus muscu
C 550	20	0.4	248417	2	AC130165	C 623	19	0.3	796	9	HS4340034	HS4340034	Homo sapi
C 551	20	0.4	248938	2	AC095006	C 624	19	0.3	867	1	AB014978	AB014978	Pseudomon
C 552	20	0.4	249245	2	AC110695	C 625	19	0.3	900	6	141261	141261	Sequence
C 553	20	0.4	250335	2	AC114441	C 626	19	0.3	900	6	161439	161439	Sequence
C 554	20	0.4	252110	2	AC123454	C 627	19	0.3	900	6	172547	172547	Sequence
C 555	20	0.4	252240	2	AC094353	C 628	19	0.3	900	6	AB056724	AB056724	Rattus no
C 556	20	0.4	252411	2	AC097717	C 629	19	0.3	1157	8	EIN18591	EIN18591	Rhesus
C 557	20	0.4	252839	2	AC095096	C 630	19	0.3	1170	8	CDA379036	CDA379036	Cynomon d
C 558	20	0.4	253874	2	AC131560	C 631	19	0.3	1188	8	ETB318590	ETB318590	Eragrosti
C 559	20	0.4	254843	2	AC095874	C 632	19	0.3	1188	8	WRB318585	WRB318585	Melinos
C 560	20	0.4	255483	2	AC095874	C 633	19	0.3	1188	8	FGB318584	FGB318584	Paspali
C 561	20	0.4	257210	2	AC098622	C 634	19	0.3	1188	8	PM318586	PM318586	Penicill m
C 562	20	0.4	257398	2	AC131885	C 635	19	0.3	1236	6	AR320776	AR320776	Sequence
C 563	20	0.4	259807	2	AC131559	C 636	19	0.3	1236	6	AX489396	AX489396	Sequence
C 564	20	0.4	261065	2	AC128931	C 637	19	0.3	1340	8	AK121530	AK121530	Oryza sat
C 565	20	0.4	261224	10	AL645527	C 638	19	0.3	1370	3	AK115471	AK115471	Ciona int
C 566	20	0.4	263259	2	AC121737	C 639	19	0.3	1433	10	AF182168	AF182168	Rattus no
C 567	20	0.4	263297	2	AC134030	C 640	19	0.3	1467	6	AX698013	AX698013	Sequence
C 568	20	0.4	263776	2	AC087335	C 641	19	0.3	1477	1	EIN3981	EIN3981	Enterobac
C 569	20	0.4	265136	2	AC094847	C 642	19	0.3	1515	4	SSIGFBP112	SSIGFBP112	Sus scrofa
C 570	20	0.4	267462	2	AC131806	C 643	19	0.3	1562	10	BC025092	BC025092	Mus muscu
C 571	20	0.4	268050	1	AL637286	C 644	19	0.3	1617	10	BC018162	BC018162	Mus muscu
C 572	20	0.4	270105	1	AC099650	C 645	19	0.3	1689	10	BC034240	BC034240	Mus muscu
C 573	20	0.4	270668	2	AC096354	C 646	19	0.3	1749	5	AB045224	AB045224	Gallus ga
C 574	20	0.4	272677	2	AC097289	C 647	19	0.3	1763	10	MM0277454	MM0277454	Mus Muscu
C 575	20	0.4	274703	2	AC110414	C 648	19	0.3	1886	9	AF131827	AF131827	Homo sapi
C 576	20	0.4	275507	2	AC099381	C 649	19	0.3	1906	10	BC002031	BC002031	Mus muscu

C 650	19	0.3	1955	9	HSU04209	U04209 Human aassoc	C 723	19	0.3	14735	10	AF317901	AF317901 Mus muscu
C 651	19	0.3	1963	3	SGNAA1	X55439 S. gregaria	C 724	19	0.3	17042	1	AE006932	AE006932 Mycobacte
C 652	19	0.3	1966	9	BC023557	BC023557 Homo sapi	C 725	19	0.3	18510	9	U72787	U72787 Homo sapien
C 653	19	0.3	1992	4	RABTCBCAI	M45777 Rabbit T-ce	C 726	19	0.3	20300	6	AX451337	AX451337 Sequence
C 654	19	0.3	2000	6	AX656080	AX656080 Sequence	C 727	19	0.3	22996	9	AC087175	AC087175 Homo sapi
C 655	19	0.3	2002	10	BC033315	BC033315 Mus muscu	C 728	19	0.3	26919	9	AP000335	AP000335 Homo sapi
C 656	19	0.3	2052	9	BC050742	BC050742 Homo sapi	C 729	19	0.3	27555	9	AC012903	AC012903 Homo sapi
C 657	19	0.3	2072	3	AB104606	AB104606 Drosophi	C 730	19	0.3	28235	2	AC014089	AC014089 Drosophi
C 658	19	0.3	2170	10	AB100266	AB100266 Mus muscu	C 731	19	0.3	30310	6	AR271852	AR271852 Sequence
C 659	19	0.3	2238	4	RABNANUCCT	M44020 Oryctolagus	C 732	19	0.3	34010	1	RSENR234	RSENR234 Sequence
C 660	19	0.3	2238	6	I11726	I11726 Sequence 1	C 733	19	0.3	38192	9	BX530088	BX530088 Human DNA
C 661	19	0.3	2304	10	RAT3AHPD	M64393 Rat 3-alpha	C 734	19	0.3	39730	1	MTVC8D5	MTVC8D5 Mycobacte
C 662	19	0.3	2345	10	AF077765	AF077765 Mus muscu	C 735	19	0.3	40806	9	MSGV140	MSGV140 Homo sapi
C 663	19	0.3	2350	9	HUMPTLL1	M4348 Human parat	C 736	19	0.3	43385	2	BX537148	BX537148 Human DNA
C 664	19	0.3	2404	9	AK025131	AK025131 Homo sapi	C 737	19	0.3	43557	2	AC147080	AC147080 Homo sapi
C 665	19	0.3	2414	6	BD127385	BD127385 Primer fo	C 738	19	0.3	43567	2	AC102848	AC102848 Homo sapi
C 666	19	0.3	2414	9	AK074960	AK074960 Homo sapi	C 739	19	0.3	45276	9	ALU12654	ALU12654 Human DNA
C 667	19	0.3	2444	6	AX713537	AX713537 Sequence	C 740	19	0.3	46651	9	AC000404	AC000404 Genomic s
C 668	19	0.3	2444	9	AK055075	AK055075 Homo sapi	C 741	19	0.3	47782	2	AC068445	AC068445 Homo sapi
C 669	19	0.3	2466	10	RAT33AD	D17310 Rattus norv	C 742	19	0.3	49850	2	AC139554	AC139554 Homo sapi
C 670	19	0.3	2477	6	AR410736	AR410736 Sequence	C 743	19	0.3	51705	6	AX695614	AX695614 Sequence
C 671	19	0.3	2477	6	AX464198	AX464198 Sequence	C 744	19	0.3	53242	9	AL365494	AL365494 Human DNA
C 672	19	0.3	2477	6	AX697578	AX697578 Sequence	C 745	19	0.3	55818	9	AC112228	AC112228 Homo sapi
C 673	19	0.3	2477	6	BD075807	BD075807 Secretary	C 746	19	0.3	58198	9	AC073524	AC073524 Homo sapi
C 674	19	0.3	2477	6	BD172367	BD172367 Secretary	C 747	19	0.3	60190	8	AP004786	AP004786 Oryza sat
C 675	19	0.3	2477	6	BD172686	BD172686 Secreterd	C 748	19	0.3	60190	6	AX697977	AX697977 Sequence
C 676	19	0.3	2477	6	BD173005	BD173005 Secreterd	C 749	19	0.3	62370	2	AL451008	AL451008 Homo sapi
C 677	19	0.3	2477	6	BD173324	BD173324 Secreterd	C 750	19	0.3	62377	2	AL360009	AL360009 Homo sapi
C 678	19	0.3	2477	6	BD173358	BD173358 Secretary	C 751	19	0.3	62380	10	AL928702	AL928702 Mouse DNA
C 679	19	0.3	2477	6	AX358496	AX358496 Homo sapi	C 752	19	0.3	63229	2	AC111184	AC111184 Homo sapi
C 680	19	0.3	2530	9	AX127587	AX127587 Homo sapi	C 753	19	0.3	63525	2	AC108463	AC108463 Homo sapi
C 681	19	0.3	2764	6	AX171368	AX171368 Sequence	C 754	19	0.3	63598	2	OS1G00045	OS1G00045 Homo sapi
C 682	19	0.3	2886	8	AX251482	AX251482 Echlinochl	C 755	19	0.3	64513	2	AC101466	AC101466 Mus muscu
C 683	19	0.3	2886	8	AX833475	AX833475 Sequence	C 756	19	0.3	64513	2	AC101466	AC101466 Mus muscu
C 684	19	0.3	2942	6	AK095247	AK095247 Homo sapi	C 757	19	0.3	64824	2	AC067782	AC067782 Homo sapi
C 685	19	0.3	2942	6	AK095247	AK095247 Homo sapi	C 758	19	0.3	65796	2	AC124642	AC124642 Mus muscu
C 686	19	0.3	2997	6	AX713389	AX713389 Sequence	C 759	19	0.3	67211	2	AC133284	AC133284 Mus muscu
C 687	19	0.3	2997	6	AX054713	AX054713 Homo sapi	C 760	19	0.3	67718	2	AC135349	AC135349 Homo sapi
C 688	19	0.3	3458	6	AX714243	AX714243 Sequence	C 761	19	0.3	67718	2	AC135349	AC135349 Homo sapi
C 689	19	0.3	3458	6	AK056454	AK056454 Homo sapi	C 762	19	0.3	68010	2	AC101197	AC101197 Mus muscu
C 690	19	0.3	3671	6	AX074078	AX074078 Sequence	C 763	19	0.3	68052	2	AC100691	AC100691 Homo sapi
C 691	19	0.3	3734	9	HSMB07851	AB047705 Homo sapi	C 764	19	0.3	68417	2	AC040891	AC040891 Homo sapi
C 692	19	0.3	3761	9	AB094095	AB094095 Homo sapi	C 765	19	0.3	69333	2	AC109778	AC109778 Homo sapi
C 693	19	0.3	3844	10	MMNADPTRH	Z49204 M. musculus	C 766	19	0.3	69586	2	AC115076	AC115076 Mus muscu
C 694	19	0.3	4121	10	AK123006	AK123006 Mus muscu	C 767	19	0.3	70089	2	AC027491	AC027491 Homo sapi
C 695	19	0.3	4142	10	BC008518	BC008518 Mus muscu	C 768	19	0.3	70574	2	AC136312	AC136312 Homo sapi
C 696	19	0.3	4474	6	I58526	I58526 Sequence 9	C 769	19	0.3	71517	9	BS000667	BS000667 Homo sapi
C 697	19	0.3	4474	6	I87001	I87001 Sequence 9	C 770	19	0.3	74098	2	AC107382	AC107382 Homo sapi
C 698	19	0.3	4474	6	AX036110	AX036110 Sequence	C 771	19	0.3	74530	2	AC108315	AC108315 Pan trogl
C 699	19	0.3	4784	6	AF249295	AF249295 Mus muscu	C 772	19	0.3	74687	10	AL683882	AL683882 Mouse DNA
C 700	19	0.3	4809	5	AP254792	AP254792 Dario rex	C 773	19	0.3	77155	9	AC117478	AC117478 Homo sapi
C 701	19	0.3	5050	9	HMDCD43	M61827 Human leuko	C 774	19	0.3	80627	2	AC027381	AC027381 Mus muscu
C 702	19	0.3	5295	10	BC060277	BC060277 Mus muscu	C 775	19	0.3	80656	2	AC110616	AC110616 Homo sapi
C 703	19	0.3	5528	10	AK123377	AK123377 Mus muscu	C 776	19	0.3	81105	2	AC109046	AC109046 Mus muscu
C 704	19	0.3	5693	10	BC036727	BC036727 Mus muscu	C 777	19	0.3	82179	2	AC100650	AC100650 Homo sapi
C 705	19	0.3	5810	10	AB011534	AB011534 Rattus no	C 778	19	0.3	82414	2	AC129265	AC129265 Homo sapi
C 706	19	0.3	5939	9	HSMB06877	AB040787 Homo sapi	C 779	19	0.3	83433	2	AC137046	AC137046 Homo sapi
C 707	19	0.3	6503	9	HSSTAL	X52075 Human gene	C 780	19	0.3	83989	9	AL603713	AL603713 Human DNA
C 708	19	0.3	6647	9	HSMB08937	BX449786 Homo sapi	C 781	19	0.3	84364	8	HS52237	HS52237 Homo sapi
C 709	19	0.3	7161	9	AB011168	AB011168 Homo sapi	C 782	19	0.3	85795	8	AP006090	AP006090 Lotus cor
C 710	19	0.3	7174	9	HSMB04580	AB011168 Homo sapi	C 783	19	0.3	86451	2	AC107037	AC107037 Homo sapi
C 711	19	0.3	8388	9	HS17CTNS1	AF112441 Homo sapi	C 784	19	0.3	88459	2	AP001136	AP001136 Homo sapi
C 712	19	0.3	9942	10	AF329470	AF329470 Rattus no	C 785	19	0.3	88712	2	AC097315	AC097315 Mus muscu
C 713	19	0.3	10029	1	AE010387	AE010387 Methanopy	C 786	19	0.3	88823	2	AL171365	AL171365 Human DNA
C 714	19	0.3	10029	1	AE014485	AE014485 Brucella	C 787	19	0.3	90100	2	AC104180	AC104180 Leishmani
C 715	19	0.3	10708	1	AE000099	AE000099 Rhizobium	C 788	19	0.3	92799	2	AC135817	AC135817 Rattus no
C 716	19	0.3	10708	10	RNTRPR2R	XC1677 Rat TRP2 g	C 789	19	0.3	92956	2	AC138707	AC138707 Homo sapi
C 717	19	0.3	11614	1	AE009582	AE009582 Brucella	C 790	19	0.3	93280	2	AL159174	AL159174 Human DNA
C 718	19	0.3	11681	1	AE009454	AE009454 Brucella	C 791	19	0.3	93280	2	AL159174	AL159174 Human DNA
C 719	19	0.3	12081	10	AL772229	AL772229 Mouse DNA	C 792	19	0.3	93393	4	AC104463	AC104463 Sus scrofa
C 720	19	0.3	12631	1	AE014357	AE014357 Brucella	C 793	19	0.3	95108	2	AC127818	AC127818 Rattus no
C 721	19	0.3	12938	1	AE012470	AE012470 Xanthomon	C 794	19	0.3	95347	9	HS0181122	HS0181122 Homo sapi
C 722	19	0.3	14207	10	AF463762	AF463762 Mus muscu	C 795	19	0.3	95546	9	AC022816	AC022816 Homo sapi

C 796	19	0.3	96009	2	EX545853	Mus muscu	869	19	0.3	142114	8	AC073166	AC073166	Oryza sat
C 797	19	0.3	99342	9	HS201D17	Human DNA	870	19	0.3	142560	2	AC021906	AC021906	Homo sapi
C 798	19	0.3	100000	9	AP000215	Human sapi	871	19	0.3	142777	10	AC006404	AC006404	Mus muscu
C 799	19	0.3	104309	9	AL391358	Human DNA	872	19	0.3	142781	2	AC141934	AC141934	Rattus no
C 800	19	0.3	104663	10	AL732560	Mouse DNA	873	19	0.3	143165	2	AL591114	AL591114	Homo sapi
C 801	19	0.3	105616	10	AL935154	Mouse DNA	874	19	0.3	143378	5	AC140945	AC140945	Gallus ga
C 802	19	0.3	107226	9	AC010618	Homo sapi	875	19	0.3	143805	2	HSB135E1	HSB135E1	Human DNA
C 803	19	0.3	107413	4	AL773560	Human sapi	876	19	0.3	144500	2	AC010731	AC010731	Homo sapi
C 804	19	0.3	107611	9	AC019221	Human sapi	877	19	0.3	144794	2	AP001182	AP001182	Homo sapi
C 805	19	0.3	108000	9	AP001180	Human sapi	878	19	0.3	145879	10	AC128915	AC128915	Mouse DNA
C 806	19	0.3	109347	9	AL137157	Human DNA	879	19	0.3	145918	3	AC012307	AC012307	Homo sapi
C 807	19	0.3	110000	2	AC009579	Continuation (2 of	880	19	0.3	146153	2	AC007452	AC007452	Drosophi
C 808	19	0.3	110000	2	AC009579	Continuation (2 of	881	19	0.3	147555	3	AC142206	AC142206	Homo sapi
C 809	19	0.3	110000	2	AC096007	Continuation (3 of	882	19	0.3	147727	2	AC011950	AC011950	Homo sapi
C 810	19	0.3	110000	2	AC096315	Continuation (6 of	883	19	0.3	147829	2	AC068806	AC068806	Mus muscu
C 811	19	0.3	110000	2	AC109940	Rattus no	884	19	0.3	149030	2	AC005375	AC005375	Homo sapi
C 812	19	0.3	110000	2	AC110111	Continuation (2 of	885	19	0.3	149087	9	AL157826	AL157826	Human sapi
C 813	19	0.3	110000	2	AC114711	Continuation (2 of	886	19	0.3	149320	2	AC126357	AC126357	Homo sapi
C 814	19	0.3	110000	2	AC116152	Mus muscu	887	19	0.3	150206	8	OSJN00083	OSJN00083	Human DNA
C 815	19	0.3	110000	2	AC117962	Continuation (6 of	888	19	0.3	151163	9	HSB105P22	HSB105P22	Mus muscu
C 816	19	0.3	110000	2	AC118670	Mus muscu	889	19	0.3	151839	2	AC102905	AC102905	Homo sapi
C 817	19	0.3	110000	2	AC118670	Continuation (2 of	890	19	0.3	151903	2	AC024349	AC024349	Homo sapi
C 818	19	0.3	110000	2	AC130075	Continuation (3 of	891	19	0.3	152058	4	AC091756	AC091756	Homo sapi
C 819	19	0.3	110000	2	AL354832	Homo sapi	892	19	0.3	152209	2	AL591023	AL591023	Homo sapi
C 820	19	0.3	110000	2	AL954350	Homo sapi	893	19	0.3	152246	9	AC023473	AC023473	Homo sapi
C 821	19	0.3	110000	2	AL954350	Continuation (4 of	894	19	0.3	152464	2	AC109229	AC109229	Mus muscu
C 822	19	0.3	111106	5	EX653531	Continuation (4 of	895	19	0.3	152827	2	AC143347	AC143347	Homo sapi
C 823	19	0.3	111891	9	AC004797	Chicken D	896	19	0.3	153021	10	AL663042	AL663042	Mouse DNA
C 824	19	0.3	115888	9	AC002468	Human Chr	897	19	0.3	153154	2	AC117770	AC117770	Mus muscu
C 825	19	0.3	116022	9	AL355533	Homo sapi	898	19	0.3	153358	2	AC145482	AC145482	Silurana
C 826	19	0.3	116311	9	AL360181	Homo sapi	899	19	0.3	153723	9	AC074101	AC074101	Homo sapi
C 827	19	0.3	118495	9	AL590556	Human DNA	900	19	0.3	154262	2	AC127629	AC127629	Rattus no
C 828	19	0.3	118593	9	HS288E11	Human DNA s	901	19	0.3	154479	2	BX255900	BX255900	Danio rer
C 829	19	0.3	119066	2	AC026377	Mus muscu	902	19	0.3	154983	2	AC018866	AC018866	Homo sapi
C 830	19	0.3	119295	2	AC127427	Magnapor	903	19	0.3	154997	2	AC143903	AC143903	Macaca mu
C 831	19	0.3	120311	10	AC006945	Mus muscu	904	19	0.3	155021	9	AC007159	AC007159	Homo sapi
C 832	19	0.3	120613	2	AC111401	AC111401 Rattus no	905	19	0.3	155278	9	HSB1655C4	HSB1655C4	Human DNA
C 833	19	0.3	123022	9	AC090717	Homo sapi	906	19	0.3	155345	9	AC137627	AC137627	Homo sapi
C 834	19	0.3	123159	2	HSAC000384	AC000384 Homo sapi	907	19	0.3	155382	9	AC127520	AC127520	Homo sapi
C 835	19	0.3	123489	9	HSBA27F12	AC109914 Human DNA	908	19	0.3	155663	3	AC092048	AC092048	Homo sapi
C 836	19	0.3	124289	9	AC010285	AC010285 Homo sapi	909	19	0.3	155670	3	AC012097	AC012097	Drosophi
C 837	19	0.3	124392	9	AL136451	Human DNA	910	19	0.3	155684	2	AL390205	AL390205	Human sapi
C 838	19	0.3	126753	2	AF252826	Human sapi	911	19	0.3	155824	2	AC087296	AC087296	Homo sapi
C 839	19	0.3	126894	9	AL157875	Human DNA	912	19	0.3	156122	2	AC127546	AC127546	Homo sapi
C 840	19	0.3	127431	9	AC027796	AC027796 Homo sapi	913	19	0.3	156264	9	AC103521	AC103521	Homo sapi
C 841	19	0.3	127544	9	CNS01RG4	AL157756 Human chr	914	19	0.3	156339	2	AC132073	AC132073	Bos tauri
C 842	19	0.3	127677	9	AC102945	AC102945 Homo sapi	915	19	0.3	157063	10	AL928850	AL928850	Mouse DNA
C 843	19	0.3	127690	2	AC138628	AC138628 Homo sapi	916	19	0.3	157340	2	AC015652	AC015652	Homo sapi
C 844	19	0.3	128090	2	AC108572	AC108572 Pan trogl	917	19	0.3	157519	5	AC146347	AC146347	Pan trogl
C 845	19	0.3	128608	10	AL928638	AL928638 Mouse DNA	918	19	0.3	157530	9	AC146347	AC146347	Pan trogl
C 846	19	0.3	129492	2	AC138466	AC138466 Homo sapi	919	19	0.3	157631	2	AC1330421	AC1330421	Homo sapi
C 847	19	0.3	129506	2	AC136100	AC136100 Rattus no	920	19	0.3	157654	2	AC091791	AC091791	Sus scrofa
C 848	19	0.3	130023	2	AC011843	AC011843 Homo sapi	921	19	0.3	157964	2	AC134625	AC134625	Rattus no
C 849	19	0.3	131189	2	AC0101756	AC0101756 Mus muscu	922	19	0.3	158041	2	AC140688	AC140688	Homo sapi
C 850	19	0.3	131353	9	HS608115	AL021707 Human DNA	923	19	0.3	158156	2	AL357393	AL357393	Homo sapi
C 851	19	0.3	131888	9	HS11057D4	AL121777 Human DNA	924	19	0.3	158274	2	AC013592	AC013592	Homo sapi
C 852	19	0.3	132150	9	AC005586	AC005586 Homo sapi	925	19	0.3	158307	10	AL592112	AL592112	Mouse DNA
C 853	19	0.3	132336	2	AP215842	AF215842 Homo sapi	926	19	0.3	158338	2	AC024720	AC024720	Homo sapi
C 854	19	0.3	132703	8	CNS080CBT	AC145489 Oryza sat	927	19	0.3	158449	9	AC011869	AC011869	Mus muscu
C 855	19	0.3	133455	2	AC145489	AC145489 Homo sapi	928	19	0.3	158510	2	AC079050	AC079050	Homo sapi
C 856	19	0.3	133698	2	AC143858	AC143858 Macaca mu	929	19	0.3	158920	2	AC135955	AC135955	Canis fam
C 857	19	0.3	133898	2	AC112917	AC112917 Homo sapi	930	19	0.3	158976	2	AC117771	AC117771	Mus muscu
C 858	19	0.3	135839	9	AC006060	AC006060 Homo sapi	931	19	0.3	159065	3	AC091634	AC091634	Drosophi
C 859	19	0.3	135854	2	AC116945	AL035846 Human DNA	932	19	0.3	159266	2	AC026097	AC026097	Homo sapi
C 860	19	0.3	136579	2	HS998G20	AC108376 Pan trogl	933	19	0.3	159397	2	AC142212	AC142212	Mus muscu
C 861	19	0.3	136778	2	AC108376	AC108376 Rattus no	934	19	0.3	159720	2	AC068716	AC068716	Homo sapi
C 862	19	0.3	136956	2	AC137338	AC137338 Homo sapi	935	19	0.3	159902	2	AC025571	AC025571	Homo sapi
C 863	19	0.3	138242	10	AC122519	AC122519 Mus muscu	936	19	0.3	159983	2	AC013713	AC013713	Homo sapi
C 864	19	0.3	138824	8	AC051633	AC051633 Oryza sat	937	19	0.3	160074	9	AC105361	AC105361	Homo sapi
C 865	19	0.3	139376	9	AC020659	AC020659 Homo sapi	938	19	0.3	160109	9	AL161727	AL161727	Human DNA
C 866	19	0.3	139918	9	AC135564	AC135564 Mus muscu	939	19	0.3	160396	9	AC107421	AC107421	Homo sapi
C 867	19	0.3	141017	10	AC126444	Human DNA	940	19	0.3	160429	2	AC136950	AC136950	Homo sapi
C 868	19	0.3	141079	9	HSB216C14	AL121694 Human DNA	941	19	0.3	160429	2	AC136950	AC136950	Homo sapi

QY	781	CTGTAGATGGTGTCTCCACCAAGGCTCCACCAACCAAGACAGGTATATCATCAACTC	840
Db	781	CTTAGATAGTGTCTCCACAGGCTCCACCAAGAGCAGGTATATCATCAACTC	840
QY	841	GTGAGGCTTTCCAGAGGACACACAGCCTTCAACTCTATGAGAGGTGCCATTGGCTGT	900
Db	841	GTGAGGCTTTCCAGAGGACACACAGCCTTCAACTCTATGAGAGGTGCCATTGGCTGT	900
QY	901	GAGCCCAETGGGGTGGAGTACCCTGTGTCAAGCTGTCACTGTCCAAAGCGGGGCC	960
Db	901	GAGCCCAETGGGGTGGAGTACCCTGTGTCAAGCTGTCACTGTCCAAAGCGGGGCC	960
QY	961	GTGCTTGGCAGGACCTTGGAGTCCATCAGATATATGACTGCTCTTCAACCGTCTTCC	1020
Db	961	GTGCTTGGCAGGACCTTGGAGTCCATCAGATATATGACTGCTCTTCAACCGTCTTCC	1020
QY	1021	AAGGCCCAAGAGCGGAAAAATGAAATCCCTGATAGTGGCCCTGTGCATCTTCACTTG	1080
Db	1021	AAGGCCCAAGAGCGGAAAAATGAAATCCCTGATAGTGGCCCTGTGCATCTTCACTTG	1080
QY	1081	AAGCAGATPAAATGACCGGATTAAAGAGGGCGTCAGTCTTGTACCCGGGGAGAGGACG	1140
Db	1081	AAGCAGATPAAATGACCGGATTAAAGAGGGCGTCAGTCTTGTACCCGGGGAGAGGACG	1140
QY	1141	CTGCACTGTGCTGTGCTCAAGGTGAAGACATCCCTTGCACAGTGGCTCTTAAACATT	1200
Db	1141	CTGCACTGTGCTGTGCTCAAGGTGAAGACATCCCTTGCACAGTGGCTCTTAAACATT	1200
QY	1201	GACGATTAATTCTGTGGCTGTGACATGAAATGCTCCCTGGGAGGTGCCGATAGTGGCT	1260
Db	1201	GACGATTAATTCTGTGGCTGTGACATGAAATGCTCCCTGGGAGGTGCCGATAGTGGCT	1260
QY	1261	GGAATTCCTGCTTTCACGAGAGACAGGACCGCATGAGCTGTGTCATCGCATATGTCTAC	1320
Db	1261	GGAATTCCTGCTTTCACGAGAGACAGGACCGCATGAGCTGTGTCATCGCATATGTCTAC	1320
QY	1321	AAGAACCACTCTGTGGCTTTGTGGGCAACMAAGTGGCAAGCTGAAGAAATCCGGGTG	1380
Db	1321	AAGAACCACTCTGTGGCTTTGTGGGCAACMAAGTGGCAAGCTGAAGAAATCCGGGTG	1380
QY	1381	GATGAGCCAGGGGCAAGCGCCTCCAGATGAGCGGTGAGGTGTGACCCCGGCCA	1440
Db	1381	GATGAGCCAGGGGCAAGCGCCTCCAGATGAGCGGTGAGGTGTGACCCCGGCCA	1440
QY	1441	GTCTCCGGGATATGACCTTCTCCAAGACCAAGCACTCTACATCATGTCAAGAGG	1500
Db	1441	GTCTCCGGGATATGACCTTCTCCAAGACCAAGCACTCTACATCATGTCAAGAGG	1500
QY	1501	CAGCTCACAGATGCCGTGTGGAAGTCCGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT	1560
Db	1501	CAGCTCACAGATGCCGTGTGGAAGTCCGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT	1560
QY	1561	GGCTTAGGCGAACCCCACTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT	1620
Db	1561	GGCTTAGGCGAACCCCACTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT	1620
QY	1621	CGGTGTGAGCGGTCCAAAGAGCCCGCAGGTTTGCTGGAGATGAAACAATGTGTCCGG	1680
Db	1621	CGGTGTGAGCGGTCCAAAGAGCCCGCAGGTTTGCTGGAGATGAAACAATGTGTCCGG	1680
QY	1681	CTGAAGGTTCATCCCAACAATATCTCCGTCTGTCAAGTCAAGTGTCTGTGTGTGTGT	1740
Db	1681	CTGAAGGTTCATCCCAACAATATCTCCGTCTGTCAAGTCAAGTGTCTGTGTGTGTGT	1740
QY	1741	ACGTACAATGTCCCGAGCTGTCAACTGACCTGACCTGACCTTTGAGACCTGTCAAG	1800
Db	1741	ACGTACAATGTCCCGAGCTGTCAACTGACCTGACCTGACCTTTGAGACCTGTCAAG	1800
QY	1801	ATGATATGGCTGT	1860
Db	1801	ATGATATGGCTGT	1860
QY	1861	CCCGGATATCAAGAGATGGGACCAACATGTGTACAGCTTCAAGTCAAAATCAAG	1920

Db	1861	CCCCGGATTCATCAAGAGAAATGGGGACCAACATGCTCTTAAGCTTCAGCTTCAATCAAG	1920
QY	1921	GAGACCCGACATGACCTTTCGCCAGCACAGCTTTGTCTTCTPACATTCGAGCGTCCACAAT	1980
Db	1921	GAGACCCGACATGACCTTTCGCCAGCACAGCTTTGTCTTCTPACATTCGAGCGTCCACAAT	1980
QY	1981	TCCAGCCGCTGCTCCGCTGGAGAGTCCATACCGCTCCAGCTGGTGTAAATACCGGCACTGC	2040
Db	1981	TCCAGCCGCTGCTCCGCTGGAGAGTCCATACCGCTCCAGCTGGTGTAAATACCGGCACTGC	2040
QY	2041	TGCACCCATGACCCCAAGACCTGCTCTTCCAGAGAGCCGAGTGAAGCTGCCGAGAC	2100
Db	2041	TGCACCCATGACCCCAAGACCTGCTCTTCCAGAGAGCCGAGTGAAGCTGCCGAGAGAC	2100
QY	2101	TGCCCCAGCGCTGCGAGTGGACAAATCTCGGTGCGCGGTGAGAGTATCAAGCTCATC	2160
Db	2101	TGCCCCAGCGCTGCGAGTGGACAAATCTCGGTGCGCGGTGAGAGTATCAAGCTCATC	2160
QY	2161	ACGCTGAAAGCCAAAGAACTCCCCAGACCCCAAGCTGTGAGCGTGGCTACGAATGCATC	2220
Db	2161	ACGCTGAAAGCCAAAGAACTCCCCAGACCCCAAGCTGTGAGCGTGGCTACGAATGCATC	2220
QY	2221	CTCAACATTCAGGGCAGCGAGCAGCAGCGTCCGCTTCGCTTCAACAGCTCCAGCGTA	2280
Db	2221	CTCAACATTCAGGGCAGCGAGCAGCAGCGTCCGCTTCGCTTCAACAGCTCCAGCGTA	2280
QY	2281	CAGTCCCGAACAACCTCTTATTCCTATGAAAGGATGAGATCAACAACCTGCCGTGAG	2340
Db	2281	CAGTCCCGAACAACCTCTTATTCCTATGAAAGGATGAGATCAACAACCTGCCGTGAG	2340
QY	2341	TTGACAGTCTGTGAAATGGGCACTTCAACATTTGACAAACCAAGCTCAGAAATTAAGTTAC	2400
Db	2341	TTGACAGTCTGTGAAATGGGCACTTCAACATTTGACAAACCAAGCTCAGAAATTAAGTTAC	2400
QY	2401	CTTCACAAAGTGTGAGCCATGCTGAGAGCGTGGCGCTGAGCTCAAGGCTGACCCAGAC	2460
Db	2401	CTTCACAAAGTGTGAGCCATGCTGAGAGCGTGGCGCTGAGCTCAAGGCTGACCCAGAC	2460
QY	2461	TTGCGCATGTGCTGTGCTGCGACAGGCGCCAGGCGCAGTGCACTCTGCGCGACACTGCTCC	2520
Db	2461	TTGCGCATGTGCTGTGCTGCGACAGGCGCCAGGCGCAGTGGCACCTGCGCGACACTGCTCC	2520
QY	2521	CAGAGAGCCCATGTGGCTGTGAGCTGTCTGATGCCAAGCAAGTGCACAAACCCCGGATC	2580
Db	2521	CAGAGAGCCCATGTGGCTGTGAGCTGTCTGATGCCAAGCAAGTGCACAAACCCCGGATC	2580
QY	2581	ACAGAGATTAATCCCGGTGCACAGGCGCCCGGGAGAGGGGACCAAGGTCATCATTCGAGG	2640
Db	2581	ACAGAGATTAATCCCGGTGCACAGGCGCCCGGGAGAGGGGACCAAGGTCATCATTCGAGG	2640
QY	2641	GAGAACTGTGGCTGTGAAATTCGCGACATGCGCTTCCATGTCAAGGTTGCTGGCGGTGAG	2700
Db	2641	GAGAACTGTGGCTGTGAAATTCGCGACATGCGCTTCCATGTCAAGGTTGCTGGCGGTGAG	2700
QY	2701	TGCAGACCTTTAGTGGATGTTACATCCCTGSCAACAAGATCGATGTGAGATGGGGAG	2760
Db	2701	TGCAGACCTTTAGTGGATGTTACATCCCTGSCAACAAGATCGATGTGAGATGGGGAG	2760
QY	2761	GCCAAAGCCAGCCAGCATGCAAGGCTTGTGTGAGATCTTGCTGTGCTGTGTGCGCTTGA	2820
Db	2761	GCCAAAGCCAGCCAGCATGCAAGGCTTGTGTGAGATCTTGCTGTGCTGTGTGCGCTTGA	2820
QY	2821	TTCAATGGCGGCTGCTCAACGCTCTATTAATTCACTTCAAGCACTGCTCTCATGACTGAG	2880
Db	2821	TTCAATGGCGGCTGCTCAACGCTCTATTAATTCACTTCAAGCACTGCTCTCATGACTGAG	2880
QY	2881	CCCAAGCCGGGGGCCCATGTCCGAGAGGAGCCAAATGACATCAAGGACCAACCTGAT	2940
Db	2881	CCCAAGCCGGGGGCCCATGTCCGAGAGGAGCCAAATGACATCAAGGACCAACCTGAT	2940
QY	2941	GCCGGAACCAACGTGTGTGTGATGTTTGGAAAGAGCGCTGTCTTCTCAAGCGATCT	3000
Db	2941	GCCGGAACCAACGTGTGTGTGATGTTTGGAAAGAGCGCTGTCTTCTCAAGCGATCT	3000

Db 2941 GCCGGAAGCAACGTGCTGATGTTGGAAGACGCCCTGTCCTTCCACAGGCCATCT 3000
QY 3001 CCATCTCAATGTTCTGCAACCAACATCCTCAGATGAGTGCTAGAGATGAAGTGTG 3060
Db 3001 CCATCTCAATGTTCTGCAACCAACATCCTCAGATGAGTGCTAGAGATGAAGTGTG 3060
QY 3061 GTGACGTGACAGGGCCAGATCCACAGGACCTGGCTTTCAGTATGTAAGAACCC 3120
Db 3061 GTGACGTGACAGGGCCAGATCCACAGGACCTGGCTTTCAGTATGTAAGAACCC 3120
QY 3121 ACCATCTGCGGATTTGAGCCAGAAATGAGCAATTGTCAGTGAACACACCCATGCGGTA 3180
Db 3121 ACCATCTGCGGATTTGAGCCAGAAATGAGCAATTGTCAGTGAACACACCCATGCGGTA 3180
QY 3181 TGGGGAGACCCCACTGAGACCTATACAGAACCCCGAGATCCGTCGCAAGATGAGGAG 3240
Db 3181 TGGGGAGACCCCACTGAGACCTATACAGAACCCCGAGATCCGTCGCAAGATGAGGAG 3240
QY 3241 GAGCACATCAATATCTGTGAGGTTCTGAACGCTACTAGATGACCTGTCAAGCGCCGCC 3300
Db 3241 GAGCACATCAATATCTGTGAGGTTCTGAACGCTACTAGATGACCTGTCAAGCGCCGCC 3300
QY 3301 CTGCTCTGGGTCCTGACCAAGTCTGACCTGACCGAGAGGCCCGAGAGTTGGCTTC 3360
Db 3301 CTGCTCTGGGTCCTGACCAAGTCTGACCGAGAGGCCCGAGAGTTGGCTTC 3360
QY 3361 ATCTGGAACAAGTCAGTCCCTGCTCATCTCAACAAGAACCACTTCACTATATCC 3420
Db 3361 ATCTGGAACAAGTCAGTCCCTGCTCATCTCAACAAGAACCACTTCACTATATCC 3420
QY 3421 AACCGGTGTTGAGGCTTGTGTCCTCAAGAAATCTGAGCTGAAAGCTTGACGCC 3480
Db 3421 AACCGGTGTTGAGGCTTGTGTCCTCAAGAAATCTGAGCTGAAAGCTTGACGCC 3480
QY 3481 ATCATCTTAAAGGGCAAGAACTGATCCGCGCTGTGGCTGGGGCAAGTGAAGCTGAC 3540
Db 3481 ATCATCTTAAAGGGCAAGAACTGATCCGCGCTGTGGCTGGGGCAAGTGAAGCTGAC 3540
QY 3541 TACACTGTGCTGTTGGGGAAGACCGTGCACCGTGCACCGTGCATGATGTCACGTGTC 3600
Db 3541 TACACTGTGCTGTTGGGGAAGACCGTGCACCGTGCACCGTGCATGATGTCACGTGTC 3600
QY 3601 TGGGAGTCCGCCAAGCTCAAGGAGGAGACCGTGCACCGTGCATGATGTCACGTGTC 3660
Db 3601 TGGGAGTCCGCCAAGCTCAAGGAGGAGACCGTGCACCGTGCATGATGTCACGTGTC 3660
QY 3661 TACTCCCGGGAGTGGTGAATGATGCCCGGACACCGCTCAGCGTCCCGCATGCTC 3720
Db 3661 TACTCCCGGGAGTGGTGAATGATGCCCGGACACCGCTCAGCGTCCCGCATGCTC 3720
QY 3721 AGCATTCGAGTGGCTGGCGGCTCCTCATCATTTTCACTGGCGGCTGATGGCCAT 3780
Db 3721 AGCATTCGAGTGGCTGGCGGCTCCTCATCATTTTCACTGGCGGCTGATGGCCAT 3780
QY 3781 AACCGAAGTCCCGGAAAGTGAAGCTCAAGCTGAAGCGGCTGAGATGCAATGAGCAAC 3840
Db 3781 AACCGAAGTCCCGGAAAGTGAAGCTCAAGCTGAAGCGGCTGAGATGCAATGAGCAAC 3840
QY 3841 CTGAGTCCCGTGTGGCTTGAAGTGAAGAAAGCTTGGCGAGCTGACAGGACATC 3900
Db 3841 CTGAGTCCCGTGTGGCTTGAAGTGAAGAAAGCTTGGCGAGCTGACAGGACATC 3900
QY 3901 CATGAGCTGACAGTGAAGCTGATGAGACCGGATTCGTTCTGAGCTATAGAACTTAC 3960
Db 3901 CATGAGCTGACAGTGAAGCTGATGAGACCGGATTCGTTCTGAGCTATAGAACTTAC 3960
QY 3961 ACCATGCGGCTGTCTTCCAGAAATTGAAGCAACCTGTCTCCGGGACCTTGAAGTTC 4020
Db 3961 ACCATGCGGCTGTCTTCCAGAAATTGAAGCAACCTGTCTCCGGGACCTTGAAGTTC 4020
QY 4021 CCGGCTACCGGACGAGCTGTGAGAAAGGCTGAAGCTTTGCGCCAGCTCATCAAC 4080
Db 4021 CCGGCTACCGGACGAGCTGTGAGAAAGGCTTTGCGCCAGCTCATCAAC 4080
QY 4081 AACAGGTCCTGCTGCTCTTCAATCCGACCGCTTGAAGTCCGAGGATGCTCCATG 4140
Db 4081 AACAGGTCCTGCTGCTCTTCAATCCGACCGCTTGAAGTCCGAGGATGCTCCATG 4140
QY 4141 CCGGACCGTGGCAAGTGGCTCACTCATCATGACCGTGTGCAAGAGCAAGCTGATAC 4200
Db 4141 CCGGACCGTGGCAAGTGGCTCACTCATCATGACCGTGTGCAAGAGCAAGCTGATAC 4200
QY 4201 GCGACTGATGCTGAGAGAGTGTGTCGCGGACCTCATTTGACAGAAACCTGAGAGCAG 4260
Db 4201 GCGACTGATGCTGAGAGAGTGTGTCGCGGACCTCATTTGACAGAAACCTGAGAGCAG 4260
QY 4261 GCCACTGATGCTGAGAGAGTGTGTCGCGGACCTCATTTGACAGAAACCTGAGAGCAG 4320
Db 4261 AACCAAGCTGAGTGTGCTGAGAGAGTGTGTCGCGGACCTCATTTGACAGAAACCTGAGAGCAG 4320
QY 4321 TGGTTTACTTTCCTCTCTTCAAGTTCTTCAAGAGAGTGTGTCGCGGACCTCTTCTCC 4380
Db 4321 TGGTTTACTTTCCTCTCTTCAAGTTCTTCAAGAGAGTGTGTCGCGGACCTCTTCTCC 4380
QY 4381 TGGTTTACTTTCCTCTCTTCAAGTTCTTCAAGAGAGTGTGTCGCGGACCTCTTCTCC 4440
Db 4381 TGGTTTACTTTCCTCTCTTCAAGTTCTTCAAGAGAGTGTGTCGCGGACCTCTTCTCC 4440
QY 4441 GCGCGTACTCTTGAAGGAGCAAGTCTATCCGCGAGAGATGACTTCAAAACCTTG 4500
Db 4441 GCGCGTACTCTTGAAGGAGCAAGTCTATCCGCGAGAGATGACTTCAAAACCTTG 4500
QY 4501 GTCCTGAGTGTGAGAGGAGCAAGTCTATCCGCGAGAGATGACTTCAAAACCTTG 4560
Db 4501 GTCCTGAGTGTGAGAGGAGCAAGTCTATCCGCGAGAGATGACTTCAAAACCTTG 4560
QY 4561 AACCTGACACCATCACTAGGTCAGAGAAAGATTTGATGTCATCTTCAAGATGTG 4620
Db 4561 AACCTGACACCATCACTAGGTCAGAGAAAGATTTGATGTCATCTTCAAGATGTG 4620
QY 4621 CCTGTCCCAACCGGCCAAAGCTGAGATGAGATGAGTGTGAGAGTGGCAAGAGAGTGG 4680
Db 4621 CCTGTCCCAACCGGCCAAAGCTGAGATGAGATGAGTGTGAGAGTGGCAAGAGAGTGG 4680
QY 4681 GCAAGGATGATCTTGAAGATGAAGACATCAACCAAGATTGAGATGTAAGCGA 4740
Db 4681 GCAAGGATGATCTTGAAGATGAAGACATCAACCAAGATTGAGATGTAAGCGA 4740
QY 4741 CTGAAACACTGGCCCACTACAGAGTGCAGATGCTTCCGTGGGCAATTAAGTCCAG 4800
Db 4741 CTGAAACACTGGCCCACTACAGAGTGCAGATGCTTCCGTGGGCAATTAAGTCCAG 4800
QY 4801 CAGGTGACAGCTATPACAGAGTGAACAACCTCAAGCTTCCAGAGCTCAAGCAAGTAA 4860
Db 4801 CAGGTGACAGCTATPACAGAGTGAACAACCTCAAGCTTCCAGAGCTCAAGCAAGTAA 4860
QY 4861 TATGAAAACATGATCCGATACACGCGGACCGGCAAGCTTCCAGAGCTCAAGCAAGTAA 4920
Db 4861 TATGAAAACATGATCCGATACACGCGGACCGGCAAGCTTCCAGAGCTCAAGCAAGTAA 4920
QY 4921 ATCACTCTGACCTGAGAGTGAAGTCAAGATGTGAGCACTTGAAGAACCAAGAGCAC 4980
Db 4921 ATCACTCTGACCTGAGAGTGAAGTCAAGATGTGAGCACTTGAAGAACCAAGAGCAC 4980
QY 4981 GAGACCAAGAGAGGAGGAGACCGGGGAGCAAGATGTGTCTGAAATCTAACCTGAG 5040
Db 4981 GAGACCAAGAGAGGAGGAGACCGGGGAGCAAGATGTGTCTGAAATCTAACCTGAG 5040
QY 5041 CTCCTGACCAATPAGGACACCTGCAAGATTTGATGACCTGTTGAGACATCTTC 5100
Db 5041 CTCCTGACCAATPAGGACACCTGCAAGATTTGATGACCTGTTGAGACATCTTC 5100
QY 5101 AGCAGGCAACCGGTGGCTGTGCGCTGCGCCCTGAGCAATGAATGATCATGCTTCCG 5160
Db 5101 AGCAGGCAACCGGTGGCTGTGCGCTGCGCCCTGAGCAATGAATGATCATGCTTCCG 5160

QY 5161 GATGAGCAGGCTGATTAACATGGCATTGATGACCCCGACGTCGCCATCTGGAAGAC 5220
DB 5161 GATGAGCAGGCTGATTAACATGGCATTGATGACCCCGACGTCGCCATCTGGAAGAC 5220
QY 5221 AATTGCTGCCCCCTGAGCTTTGGGTCAACATGATGATGAGAACCCGGAGTTGGTTGAC 5280
DB 5221 AATTGCTGCCCCCTGAGCTTTGGGTCAACATGATGATGAGAACCCGGAGTTGGTTGAC 5280
QY 5281 ATCCATTAAGAACACATCAACAGACGCTGCTCTGTGTGTGCTCAGACCTTCATGAC 5340
DB 5281 ATCCATTAAGAACACATCAACAGACGCTGCTCTGTGTGTGCTCAGACCTTCATGAC 5340
QY 5341 TCTTGTCTCCAGTGAAGACACCGGCTGGGCAAGACCTCGCCCTCCAAACAGCTGTGTAT 5400
DB 5341 TCTTGTCTCCAGTGAAGACACCGGCTGGGCAAGACCTCGCCCTCCAAACAGCTGTGTAT 5400
QY 5401 GCCAAGACATCCCAAGCTCAACAGATTTGGGTGAGAGATTACTCAGACATAGGGAAG 5460
DB 5401 GCCAAGACATCCCAAGCTCAACAGATTTGGGTGAGAGATTACTCAGACATAGGGAAG 5460
QY 5461 ATGCCAGCCATCAGCGACCAAGACATGACGATACCTGTGCTGACGATCCCGGATGAC 5520
DB 5461 ATGCCAGCCATCAGCGACCAAGACATGACGATACCTGTGCTGACGATCCCGGATGAC 5520
QY 5521 ATGAATGAGTTCAACACATGATGATGATGATGATGATGATGATGATGATGATGATGAT 5580
DB 5521 ATGAATGAGTTCAACACATGATGATGATGATGATGATGATGATGATGATGATGATGAT 5580
QY 5581 AGCAGAGAGATCCTTGAACCTCTGACACCAAGATGACGATGATGATGATGATGATGAT 5640
DB 5581 AGCAGAGAGATCCTTGAACCTCTGACACCAAGATGACGATGATGATGATGATGATGAT 5640
QY 5641 TACAACTAGAACAAAGTCAATTAACCTCATGAGCTTAGACAGCTGAATTA 5691
DB 5641 TACAACTAGAACAAAGTCAATTAACCTCATGAGCTTAGACAGCTGAATTA 5691

RESULT 2
AX492982 6367 bp DNA linear PAT 26-SEP-2002
LOCUS Sequence 19 from Patent WO02059312.
DEFINITION AX492982
ACCESSION AX492982.1 GI:23338637
VERSION
KEYWORDS
SOURCE
ORGANISM
Homo sapiens (human)
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

REFERENCE
AUTHORS
1 Kallik, D.A., Lee, S., Xu, Y., Yao, M.G., Yue, H., Bandman, O.B.,
Buford, N., Gandhi, A.R., Graul, R.C., Lai, P.G., Lu, D.A., Lu, Y.,
Tang, T.Y., Duggan, B.M., Gietzen, R.J., Hillman, J.L., Honchell, C.D.,
Ramkumar, J., Walla, N.K. and Warren, B.A.
Cell adhesion proteins
Patent: WO 02059312-A 19 01-ATG-2002;
JOURNAL INCYTE GENOMICS INC (US)
FEATURES
source
1. 6367
/organism="Homo sapiens"
/mol_type="unassigned DNA"
/db_xref="taxon:9606"
/note="Incyle ID No: 7156379CB1."

ORIGIN
Query Match 97.2%; Score 5532; DB 6; Length 6367;
Best Local Similarity 99.9%; Pred. No. 0;
Matches 5682; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 ATGAAGCAGTCCCTGGAATGACGATGATGATGATGATGATGATGATGATGATGATGATGAT 60
DB 566 ATGAAGCAGTCCCTGGAATGACGATGATGATGATGATGATGATGATGATGATGATGATGAT 625
QY 61 GGCCTCTCACTTGTCTACCCGCAAGCCCGCTGTCCAGAGACGCGTCAATT 120

DB 626 GGCCTCTCACTTGTCTACCCGCAAGCCCGCTGTCCAGAGACGCGTCAATT 685
QY 121 GTCACTTCCAGAGAGACCCCGAGGGTTTCAATCACTGTGTGTGATGAGAGACA 180
DB 686 GTCACTTCCAGAGAGACCCCGAGGGTTTCAATCACTGTGTGTGATGAGAGACA 745
QY 181 GGAACATTTACTTGGGGGCGGTCAATGGAATTTACAACTCTCAGAGACTGAAGTTC 240
DB 746 GGAACATTTACTTGGGGGCGGTCAATGGAATTTACAACTCTCAGAGACTGAAGTTC 805
QY 241 TTGTGAGCATGAGACAGAGCCGAGACAGAGACAACCCAAAGTGTACCCACCCGATC 300
DB 806 TTGTGAGCATGAGACAGAGCCGAGACAGAGACAACCCAAAGTGTACCCACCCGATC 865
QY 301 GTCCAGACCTGCAATGACCCCTTGAACCAACCAACCAATGTCAACCAATGTCTCTCA 360
DB 866 GTCCAGACCTGCAATGACCCCTTGAACCAACCAACCAATGTCAACCAATGTCTCTCA 925
QY 361 GACTACAGAGAGAACAGGCTGATGCTGTGGAGAGCTGTACCAAGCATCTGACGCTG 420
DB 926 GACTACAGAGAGAACAGGCTGATGCTGTGGAGAGCTGTACCAAGCATCTGACGCTG 985
QY 421 CTGAGGCTGAGAGACCTTTCAAGCTGGGGAGACCTTATCAATAGAGAGACATATCTG 480
DB 986 CTGAGGCTGAGAGACCTTTCAAGCTGGGGAGACCTTATCAATAGAGAGACATATCTG 1045
QY 481 TCAGGTCACAGAGAGAGGAGCTGATGCTGTGGAGAGCTGTACCAAGCATCTGAT 540
DB 1046 TCAGGTCACAGAGAGAGGAGCTGATGCTGTGGAGAGCTGTACCAAGCATCTGAT 1105
QY 541 GACAAAGCTTTCAATGACAGGCAAGGCAAGGCAAGGCAAGGCAAGGCAAGGCAAGGCA 600
DB 1106 GACAAAGCTTTCAATGACAGGCAAGGCAAGGCAAGGCAAGGCAAGGCAAGGCAAGGCA 1165
QY 601 AGCCGGAACCTGACCAAGAACTCTGAGGGGAGTGGCAAGTTCGAGTCACTTCATGAT 660
DB 1166 AGCCGGAACCTGACCAAGAACTCTGAGGGGAGTGGCAAGTTCGAGTCACTTCATGAT 1225
QY 661 GAGTTCGTGGCTCGATGATTAAGATCCCTTGGACACCTTCACATCATCTGATCTT 720
DB 1226 GAGTTCGTGGCTCGATGATTAAGATCCCTTGGACACCTTCACATCATCTGATCTT 1285
QY 721 GATATCTACTATGCTATGCTTTTGAAGTGGCACTTGTCTACTTTTGAAGCTTCA 780
DB 1286 GATATCTACTATGCTATGCTTTTGAAGTGGCACTTGTCTACTTTTGAAGCTTCA 1345
QY 781 CCTGAGATGCTGTCTCAACAGCTTCAACCAAGAGACAGTGTATATCATCAAGCTC 840
DB 1346 CCTGAGATGCTGTCTCAACAGCTTCAACCAAGAGACAGTGTATATCATCAAGCTC 1405
QY 841 GTGAGGCTTTGCAAGAGAGACACGCTTCAACTCTATATGATGAGTGGCTTGGCTGT 900
DB 1406 GTGAGGCTTTGCAAGAGAGACACGCTTCAACTCTATATGATGAGTGGCTTGGCTGT 1465
QY 901 GAGCGAGTGGGTGAGTACCGCTGCTGAGGCTGAGCTGCTTCAAAAGCGGGGCGC 960
DB 1466 GAGCGAGTGGGTGAGTACCGCTGCTGAGGCTGAGCTGCTTCAAAAGCGGGGCGC 1525
QY 961 GTGCTTGGAGAGACCTTGGAGTCCATGATGATGATGATGATGATGATGATGATGATGAT 1020
DB 1526 GTGCTTGGAGAGACCTTGGAGTCCATGATGATGATGATGATGATGATGATGATGATGAT 1585
QY 1021 AAGGGCCAGAGGGGAAATGAAATCCCTGAGTGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1080
DB 1586 AAGGGCCAGAGGGGAAATGAAATCCCTGAGTGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1645
QY 1081 AAGCAGATTAATGACCGCATTAAGAGAGGCGCTGAGTCTTGTACCGGGGCAAGGAGCAG 1140
DB 1646 AAGCAGATTAATGACCGCATTAAGAGAGGCGCTGAGTCTTGTACCGGGGCAAGGAGCAG 1705
QY 1141 CTGAGCTGGGCTGGCTCAAGGTGAAGACATCCCTGACAGAGTGGCTCTTAACAT 1200

Db	1706	CTGGACCTGGCCTTGGCTTCAGGTGAAAGACATCCCTCTGACAGACTGGGCTCTTAACATT	1765
Qy	1201	GACGATAACTTCTGTGGCTTGAGCATGATGCTCCCTTGAGAGTGTCCGACATGTGGCT	1260
Db	1766	GACGATAACTTCTGTGGCTTGAGCATGATGCTCCCTTGAGAGTGTCCGACATGTGGCT	1825
Qy	1261	GGAAATTCCTGCTTCACGAGAGACAGGGACCGCATGACGTCTGTCAATCGATATGCTAC	1320
Db	1826	GGAAATTCCTGCTTCACGAGAGACAGGGACCGCATGACGTCTGTCAATCGATATGCTAC	1885
Qy	1321	AAGAACCACTCTTGAGCCTTTGTGGGACCAAAAGTGGCACTGAAGAAATCCGGGTG	1380
Db	1886	AAGAACCACTCTTGAGCCTTTGTGGGACCAAAAGTGGCACTGAAGAAATCCGGGTG	1945
Qy	1381	GATGGACCCAGGGGCAAGCCCTCGAATATAGACGGTGCAGGTGTGGACCCCGGCCA	1440
Db	1946	GATGGACCCAGGGGCAAGCCCTCGAATATAGACGGTGCAGGTGTGGACCCCGGCCA	2005
Qy	1441	GTCCTCCGGATATGCTTCTCCAGGACCAAGACCACTATCATATGTACAGAGG	1500
Db	2006	GTCCTCCGGATATGCTTCTCCAGGACCAAGACCACTATCATATGTACAGAGG	2065
Qy	1501	CAGCTCACAGAGTCCCTGTGAGTCCGTGTGATCAGATCAGAGCTCGCGAGTGCCT	1560
Db	2066	CAGCTCACAGAGTCCCTGTGAGTCCGTGTGATCAGATCAGAGCTCGCGAGTGCCT	2125
Qy	1561	GGCTCAGGCGACCCCACTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT	1620
Db	2126	GGCTCAGGCGACCCCACTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT	2185
Qy	1621	CGGTGTAGCGGTCCAAAGAGCCCGCAGTTCCTCGAGATGAAAGAGTGTGTCCG	1680
Db	2186	CGGTGTAGCGGTCCAAAGAGCCCGCAGTTCCTCGAGATGAAAGAGTGTGTCCG	2245
Qy	1681	CTGACGTGTCCATCCCAATATATCTCGTCTCTCAGTACAACTGTGTGTGTGTGTGTGT	1740
Db	2246	CTGACGTGTCCATCCCAATATATCTCGTCTCTCAGTACAACTGTGTGTGTGTGTGTGT	2305
Qy	1741	ACGTACAAATGTCCCGAGCTGTCAAGTGGCTCAACTGCACTTTGAGGACTTGTCAAG	1800
Db	2306	ACGTACAAATGTCCCGAGCTGTCAAGTGGCTCAACTGCACTTTGAGGACTTGTCAAG	2365
Qy	1801	ATGTGATGGCTGT	1860
Db	2366	ATGTGATGGCTGT	2425
Qy	1861	CCCCGGATATCACAGAGATGGGACCAACATGTCTGACAGCTTCAGCTCAATCAAG	1920
Db	2426	CCCCGGATATCACAGAGATGGGACCAACATGTCTGACAGCTTCAGCTCAATCAAG	2485
Qy	1921	GAGACCGGATATCTTGTGCGACGACCAAGCTTTGTCTTCAATATGAGGTCCAAAT	1980
Db	2486	GAGACCGGATATCTTGTGCGACGACCAAGCTTTGTCTTCAATATGAGGTCCAAAT	2545
Qy	1981	TGCTGTCTGTCTTGT	2040
Db	2546	TGCTGTCTGTCTTGT	2605
Qy	2041	TGCAACCAATGACCCCAAGCTGTCTTCTCAAGAAAGCCAGATGAACTGCGCGAGAC	2100
Db	2606	TGCAACCAATGACCCCAAGCTGTCTTCTCAAGAAAGCCAGATGAACTGCGCGAGAC	2665
Qy	2101	TGCCCCAGCTCTGTGAGTGTGACAAAGTCTGTGTGTGTGTGTGTGTGTGTGTGTGTGT	2160
Db	2666	TGCCCCAGCTCTGTGAGTGTGACAAAGTCTGTGTGTGTGTGTGTGTGTGTGTGTGTGT	2725
Qy	2161	ACGCTGAAAGCAAGAACTTCCCGAGCCCAAGTCTGGGACAGCGTGTACGAAATGCATC	2220
Db	2726	ACGCTGAAAGCAAGAACTTCCCGAGCCCAAGTCTGGGACAGCGTGTGTACGAAATGCATC	2785
Qy	2221	CTCAACATTCAAGGGCAGGACAGCGAGTGTCCCGCTTGTGTGTGTGTGTGTGTGTGTGT	2280
Db	2786	CTCAACATTCAAGGGCAGGACAGCGAGTGTCCCGCTTGTGTGTGTGTGTGTGTGTGTGT	2845

QY	2281	CAGTGCAGAAACCTCTTAATTCCTATGAAGGGATGAGATCAACAACCTGCCCTGTGGAG	2340
Db	2846	CAGTGCAGAAACCTCTTAATTCCTATGAAGGGATGAGATCAACAACCTGCCCTGTGGAG	2905
QY	2341	TTGACAGTCGTGTGAATGGGCACTTCAACATTGACAAACCCAGCTCAGATTAAGTTTCA	2400
Db	2906	TTGACAGTCGTGTGAATGGGCACTTCAACATTGACAAACCCAGCTCAGATTAAGTTTCA	2965
QY	2401	CTTCAACAAGTGTGAGCCCATGCTGAGAGCTGGGGCTGTGCTCAAGGCTGACCCAGAC	2460
Db	2966	CTTCAACAAGTGTGAGCCCATGCTGAGAGCTGGGGCTGTGCTCAAGGCTGACCCAGAC	3025
QY	2461	TTTGCATGTGCTGTGTCCAGGGGCCAGGCCAGTGCACCCTGGCCGACACTGCCCTGCC	2520
Db	3026	TTTGCATGTGCTGTGTCCAGGGGCCAGGCCAGTGCACCCTGGCCGACACTGCCCTGCC	3085
QY	2521	CAGGAGAGCCAGTGGCTGGAGCTGTCTGTGTCCAAAAGCAAGTGCATAAATCCCCGATC	2580
Db	3086	CAGGAGAGCCAGTGGCTGGAGCTGTCTGTGTCCAAAAGCAAGTGCATAAATCCCCGATC	3145
QY	2581	ACAGAGATTAATCCCGGTGACAGGCCCCCGGAAAGGGGGCACAAAGTCACTATCCGAGG	2640
Db	3146	ACAGAGATTAATCCCGGTGACAGGCCCCCGGAAAGGGGGCACAAAGTCACTATCCGAGG	3205
QY	2641	GAGAACTGTGGCTTGGAAATTTCCGCACATGGCTCCATGTCAAGTTTCTGGGTGTGGAG	2700
Db	3206	GAGAACTGTGGCTTGGAAATTTCCGCACATGGCTCCATGTCAAGTTTCTGGGTGTGGAG	3265
QY	2701	TGCAGCCCTTTAGTGAATGGTTATCATCCCTGTACAAACAGATCGTGTGTGATGTGGGGAG	2760
Db	3266	TGCAGCCCTTTAGTGAATGGTTATCATCCCTGTACAAACAGATCGTGTGTGATGTGGGGAG	3325
QY	2761	GCCAAAGCCAGCCAGCATGACAGGCTTGGTGGAGATCTTGCTGGCTGTGTGTCCGGCTTGA	2820
Db	3326	GCCAAAGCCAGCCAGCATGACAGGCTTGGTGGAGATCTTGCTGGCTGTGTGTCCGGCTTGA	3385
QY	2821	TTTATGGCCGGGTCTCTCAAGCTCTATTAATCTTCAACACTGACCTCTCAGATCTTAAAG	2880
Db	3386	TTTATGGCCGGGTCTCTCAAGCTCTATTAATCTTCAACACTGACCTCTCAGATCTTAAAG	3445
QY	2881	CCAGCCGGGGGGCCCATATCTCCGAGGGGACCCAAATGACCATCAAGGCACTAACCTGAAT	2940
Db	3446	CCAGCCGGGGGGCCCATATCTCCGAGGGGACCCAAATGACCATCAAGGCACTAACCTGAAT	3505
QY	2941	GCCGGAAGCAACGTGTGTGTATGTGTTGAAAGCAGCCCTGTCTCTTCCACAGCGCATCT	3000
Db	3506	GCCGGAAGCAACGTGTGTGTATGTGTTGAAAGCAGCCCTGTCTCTTCCACAGCGCATCT	3565
QY	3001	CCATCTCTAATTGTCTGCAACCAACATCTCAATGAGTGTCTAAGATGAAGTGTGTGC	3060
Db	3566	CCATCTCTAATTGTCTGCAACCAACATCTCAATGAGTGTGTCTAAGATGAAGTGTGTGC	3625
QY	3061	GTGCAAGTGTGACAGGGCCAAAGTCAACAGAGACCTTGCATTAGTATGTGAAAGACCC	3120
Db	3626	GTGCAAGTGTGACAGGGCCAAAGTCAACAGAGACCTTGCATTAGTATGTGAAAGACCC	3685
QY	3121	ACCATCTGTGGATTGAGCCAGAAATGAGCAATTGTCAETGAAAACAACCATGTGCCGTA	3180
Db	3686	ACCATCTGTGGATTGAGCCAGAAATGAGCAATTGTCAETGAAAACAACCATGTGCCGTA	3745
QY	3181	TGGGGGACCAACCTGGACCTCATACAGAAACCCAGATCCCTGTCCAAACATGAGGGAG	3240
Db	3746	TGGGGGACCAACCTGGACCTCATACAGAAACCCAGATCCCTGTCCAAACATGAGGGAG	3805
QY	3241	GAGCACTCATATATCTGTGAGTTTGAACCGTACTAGATGACTGTCAAGCGCCGCC	3300
Db	3806	GAGCACTCATATATCTGTGAGTTTGAACCGTACTAGATGACTGTCAAGCGCCGCC	3865
QY	3301	CTCGCTCTGGGTCTGACCAACAGCTGACCCGAGAGGGCCGAGAGTTTGGCTTC	3360
Db	3866	CTCGCTCTGGGTCTGACCAACAGCTGACCCGAGAGGGCCGAGAGTTTGGCTTC	3925

```
QY 3361 ATCTGACACAGTCGAGTCCCTGCTCATCTCAACAAGCAACTTCACTACATATCC 3420
Db 3326 ATCTGACACAGTCGAGTCCCTGCTCATCTCAACAAGCAACTTCACTACATATCC 3385
QY 3421 AACCCGCTGTTGAGGCTTGTGCTGCTCAAGAACTCTGAGGCTCAAGCTGACAGCC 3480
Db 3386 AACCCGCTGTTGAGGCTTGTGCTGCTCAAGAACTCTGAGGCTCAAGCTGACAGCC 4045
QY 3481 ATCATCTTAAAGGCAAGAACTGATCCCGCTGTGCTGAGGCAACGTGAAGCTGAAC 3540
Db 4046 ATCATCTTAAAGGCAAGAACTGATCCCGCTGTGCTGAGGCAACGTGAAGCTGAAC 4105
QY 3541 TACACTGTGCTGTTGGGAGAGACCGTGCAACGTGACCGTGTCAAGTGTCCAGCTGCTC 3600
Db 4106 TACACTGTGCTGTTGGGAGAGACCGTGCAACGTGACCGTGTCAAGTGTCCAGCTGCTC 4165
QY 3601 TGGAGTCCCCCACTCATTCGACAGGCAAAAGTGAATGGCCGCTGTGAGTGCATGAG 3660
Db 4166 TGGAGTCCCCCACTCATTCGACAGGCAAAAGTGAATGGCCGCTGTGAGTGCATGAG 4225
QY 3661 TACTTCCCGGGATGTGTATTCATTCGCCGAGCAGCCGCTGACGCTGCCCCCATGCTC 3720
Db 4226 TACTTCCCGGGATGTGTATTCATTCGCCGAGCAGCCGCTGACGCTGCCCCCATGCTC 4285
QY 3721 AGCATTCGAGTGGCTGGCGGCTCTCTCATCATTTTCATTCGAGCCGCTCATTTGCTAT 3780
Db 4286 AGCATTCGAGTGGCTGGCGGCTCTCTCATCATTTTCATTCGAGCCGCTCATTTGCTAT 4345
QY 3781 AAACGCAAGTCCCGCAAAAGTGAACCTCAACGCTGAAGCGCTGACATGACATGACAC 3840
Db 4346 AAACGCAAGTCCCGCAAAAGTGAACCTCAACGCTGAAGCGCTGACATGACATGACAC 4405
QY 3841 CTGAGATCCCGTGTGGCCCTGAGATGAGAGGACCTTTCGAGCTGACAGCGACATC 3900
Db 4406 CTGAGATCCCGTGTGGCCCTGAGATGAGAGGACCTTTCGAGCTGACAGCGACATC 4465
QY 3901 CATGAGTCAACAGTGAACCTGATGAGACCGGGAATTCGTTCTGACATATGAACCTTAC 3960
Db 4466 CATGAGTCAACAGTGAACCTGATGAGACCGGGAATTCGTTCTGACATATGAACCTTAC 4525
QY 3961 ACCATGCGGAGTGTCTTCCAGAAATGGAACAACACCTGTCCTCGGAGACCTTGAAGTGC 4020
Db 4526 ACCATGCGGAGTGTCTTCCAGAAATGGAACAACACCTGTCCTCGGAGACCTTGAAGTGC 4585
QY 4021 CCGGCTACCGGACAGAGACGTGTGAGAAAGGCTTGAAGCTTTTCGCCACCTCATCAAC 4080
Db 4586 CCGGCTACCGGACAGAGACGTGTGAGAAAGGCTTGAAGCTTTTCGCCACCTCATCAAC 4645
QY 4081 AACAAAGTGTCTGTCTCTTCAATCCGACGCTTGAAGTCCAGCGTATCTTCATG 4140
Db 4646 AACAAAGTGTCTGTCTCTTCAATCCGACGCTTGAAGTCCAGCGTATCTTCATG 4705
QY 4141 CGCGACCGTGGAAAGTGGCTCACTCATGACCGGTGACCGGTGACAGCAAGCTGAGTAC 4200
Db 4706 CGCGACCGTGGAAAGTGGCTCACTCATGACCGGTGACCGGTGACAGCAAGCTGAGTAC 4755
QY 4201 GCCACTGATGTGTGAGACAGAGTGTGGCCGACCTCATTTGACAAAGACCTGAGAGAC 4260
Db 4766 GCCACTGATGTGTGAGACAGAGTGTGGCCGACCTCATTTGACAAAGACCTGAGAGAC 4825
QY 4261 AACCAACCTAAGGCTGTCTGACAGAGACTGAGTCAAGTGTGAGAAAGTCTGACCAAT 4320
Db 4826 AACCAACCTAAGGCTGTCTGACAGAGACTGAGTCAAGTGTGAGAAAGTCTGACCAAT 4885
QY 4321 TGGTTACTTCTCTCTCAAGTTCTCTCAAGAGTGTGTGGAGACCTCTTCTTCC 4380
Db 4886 TGGTTACTTCTCTCTCAAGTTCTCTCAAGAGTGTGTGGAGACCTCTTCTTCC 4945
QY 4381 CTGTTCTGTGACATGAAGAGAGATGAGAAAGGACCCCATTTGACGACCATCAAGGAGAG 4440
Db 4946 CTGTTCTGTGACATGAAGAGAGATGAGAAAGGACCCCATTTGACGACCATCAAGGAGAG 5005
QY 4441 GCCCGTACTCTTGAAGAGAGCAAGCTCATTCGCGACAGATTTACTAATAAACCTTG 4500
Db 5006 GCCCGTACTCTTGAAGAGAGCAAGCTCATTCGCGACAGATTTACTAATAAACCTTG 5065
QY 4501 GTCTGAGCTGTGTACGCCCAAGATGCCAAGAGCCCGAGGTCCAGTAAAGATCTTC 4560
Db 5066 GTCTGAGCTGTGTACGCCCAAGATGCCAAGAGCCCGAGGTCCAGTAAAGATCTTC 5125
QY 4561 AACTGTGACCATATCACTCAAGTCAAGAGAGAAATTCGTGATGGCATTTGAAGATGTG 4620
Db 5126 AACTGTGACCATATCACTCAAGTCAAGAGAGAAATTCGTGATGGCATTTGAAGATGTG 5185
QY 4621 CTTTGTCCACCGGACCCAAAGCTGCAGATATGATCTGAGTGGCGACAAAGAGTGGG 4680
Db 5186 CTTTGTCCACCGGACCCAAAGCTGCAGATATGATCTGAGTGGCGACAAAGAGTGGG 5245
QY 4681 GCAAGATGATCTTGCAGAGATGAAGACATCAACCAAGATTGAGAAATTTGAAGGGA 4740
Db 5246 GCAAGATGATCTTGCAGAGATGAAGACATCAACCAAGATTGAGAAATTTGAAGGGA 5305
QY 4741 CTGAACACACTGGCCCACTACCAAGTGCAGATGATGCTGTGTGGCATTAATGTCCAAG 4800
Db 5306 CTGAACACACTGGCCCACTACCAAGTGCAGATGATGCTGTGTGGCATTAATGTCCAAG 5365
QY 4801 CAGTGAACGCTTATACGACGTGAACAACTTCACCGTGTCCAGGACCTTCAGCAACTTATG 4860
Db 5366 CAGTGAACGCTTATACGACGTGAACAACTTCACCGTGTCCAGGACCTTCAGCAACTTATG 5425
QY 4861 TATGAAACATGATTCGGTGAACCGGGCAGCCCGGACAGCTCCGCTCAGGACACTTATG 4920
Db 5426 TATGAAACATGATTCGGTGAACCGGGCAGCCCGGACAGCTCCGCTCAGGACACTTATG 5485
QY 4921 ATCACTCTGACCTTGAAGATGAGTCAAGATGTGCACTTATGTAAGAACAGACAC 4980
Db 5486 ATCACTCTGACCTTGAAGATGAGTCAAGATGTGCACTTATGTAAGAACAGACAC 5545
QY 4981 GGAAGACAGAGAGAGAGGAGCCGGGGAGCAAGATGAGTGTGAATCTACCTGACCGGA 5040
Db 5546 GGAAGACAGAGAGAGAGGAGCCGGGGAGCAAGATGAGTGTGAATCTACCTGACCGGA 5605
QY 5041 CTCCTGCGCACTTAAGGGAACACTGACAAATTTGTGATGACCTTCTTGAAGCAATCTTC 5100
Db 5606 CTCCTGCGCACTTAAGGGAACACTGACAAATTTGTGATGACCTTCTTGAAGCAATCTTC 5665
QY 5101 AGCAGGACACCGTGTCTTCCCTGCGCCCTGGCCATCAAGTACATGTTGAACCTTCTG 5160
Db 5666 AGCAGGACACCGTGTCTTCCCTGCGCCCTGGCCATCAAGTACATGTTGAACCTTCTG 5725
QY 5161 GATGAGCAGGCTGATTAACATGATGATGACCCGACAGTCCGACATACCTGAGAGAC 5220
Db 5726 GATGAGCAGGCTGATTAACATGATGATGACCCGACAGTCCGACATACCTGAGAGAC 5785
QY 5221 AATGCTTCCCTTGAAGTGTGTTGGTCAACATGATCAAGAACCCGCAAGTGTGTTGAC 5280
Db 5786 AATGCTTCCCTTGAAGTGTGTTGGTCAACATGATCAAGAACCCGCAAGTGTGTTGAC 5845
QY 5281 ATCCATTAAGACAGATCAACAGCCCTGCTCTGTGTGTGCTCAAGCCTTCAATGAC 5340
Db 5846 ATCCATTAAGACAGATCAACAGCCCTGCTCTGTGTGTGCTCAAGCCTTCAATGAC 5905
QY 5341 TCTTGTCCAGCTCAAGAGACCGGCTGGGCAAGACTGCGCTTCCAAAGAGTGTGTAT 5400
Db 5906 TCTTGTCCAGCTCAAGAGACCGGCTGGGCAAGACTGCGCTTCCAAAGAGTGTGTAT 5965
QY 5401 GCCAAGACATCCCGACGTTACAAGAAATGGAGAGGATTAATCTCAGACATTAAGGAG 5460
Db 5966 GCCAAGACATCCCGACGTTACAAGAAATGGAGAGGATTAATCTCAGACATTAAGGAG 6025
QY 5461 ATGCGACCATCAAGGACCAAGACATGAAGCGATACCTGTGAGAGTCCCGGATGAC 5520
Db 6026 ATGCGACCATCAAGGACCAAGACATGAAGCGATACCTGTGAGAGTCCCGGATGAC 6085
QY 5521 ATGAATGATTAACACATGAGTGAACCTCAGAGATCTTCTCATATGTGGCAAAATAC 5580
```

Accession	Sequence	Position
Dp	ATGATGAGTTCACACCATGATGAGTCTCGAAGATCTTTCCTATGAGGCAAAATAC	6145
Dp	6088	
Dy	AGGAGGAGATCCTTGAGACCTCTGGACACAGATGTCGGAAAGCAAACTGGCC	5640
Dy	5581	
Dp	AGGAGGAGATCCTTGAGACCTCTGGACACAGATGTCGGAAAGCAAACTGGCC	6205
Dp	6146	
Dy	TGCAACTGGAACAAGTGTATAACCTCATGAGCTTATACAGCTGA	5685
Dy	5641	
Dp	TACAACTGGAACAAGTGTATAACCTCATGAGCTTATACAGCTGA	6250
Dp	6206	

Db	961	GA	CAGGCCCCCGGAAAGGGGGCA	CCAAAGTCACTATCCGAAGGGGAAACCTGGGCTGG	1020
QY	2658	ATTTCG	GACATGCGCTCCCATGTCAAGGTTGCTGGCGTGGAGTGCAGGCCCTTTAGTGA	2717	
Db	1021	ATTTCG	GAGATCGCCCTCCCATGTCAAGGTTGCTGGCGTGGAGTGCAGGCCCTTTAGTGA	1080	
QY	2718	TG	TTACATCCCTGCAGAACAGATCGTGTGTGAATGGGGAGGCGCAAGCCCAAGCCAGCA	2777	
Db	1081	TG	TTATACCTCCCTGCAGAACAGATCGTGTGTGAATGGGGAGGCGCAAGCCCAAGCCAGCA	1140	
QY	2778	TG	CAGGCTTGTGGAGATCTCGCGTGGCTGTGTGTGCGCTGAAATTCATGTGCGGTCTC	2837	
Db	1141	TG	CAGGCTTGTGGAGATCTCGCGTGGCTGTGTGTGCGCTGAAATTCATGTGCGGTCTC	1200	
QY	2838	AC	AGCTCTAATTACTTCATGACACTGACTCTCTCAGATCTGAAGCCCAAGCCGGGGCCCAT	2897	
Db	1201	AC	AGCTCTAATTACTTCATGACACTGACTCTCTCAGATCTGAAGCCCAAGCCGGGGCCCAT	1260	
QY	2898	GT	CCGAGAGGAGCCCAAGTGAACATCAAGGCAACCACTGAATGCCGAGACACGTGCT	2957	
Db	1261	GT	CCGAGAGGAGCCCAAGTGAACATCAAGGCAACCACTGAATGCCGAGACACGTGCT	1320	
QY	2958	GAT	GATGTTTGAAGAGAGCCCTGTCTCTTCAAGGAGATCTCCATCTCATTTGCTGTG	3017	
Db	1321	GAT	GATGTTTGAAGAGAGCCCTGTCTCTTCAAGGAGATCTCCATCTCATTTGCTGTG	1380	
QY	3018	CA	ACACCATCTTCAGATGAGTGTCTAGATGAAGTGTTCGGTTCAGTGTGACAGAGC	3077	
Db	1381	CA	ACACCATCTTCAGATGAGTGTCTAGATGAAGTGTTCGGTTCAGTGTGACAGAGC	1440	
QY	3078	CA	AGTCCACACAGGACCTGTCTTCAAGTATGTGAAGACCCCAACCTGTGCGGATTTGA	3137	
Db	1441	CA	AGTCCACACAGGACCTGTCTTCAAGTATGTGAAGACCCCAACCTGTGCGGATTTGA	1500	
QY	3138	GCC	GAATGAGACATTTGTCAGTGAACACACCCATTCGCTATGTGGGGACCCACCTTGA	3197	
Db	1501	GCC	GAATGAGACATTTGTCAGTGAACACACCCATTCGCTATGTGGGGACCCACCTTGA	1560	
QY	3198	CCT	CATACAGAACCCCGACATTCGCGCCCAACGATGGAGGAAGGACATCAATATCTG	3257	
Db	1561	CCT	CATACAGAACCCCGACATTCGCGCCCAACGATGGAGGAAGGACATCAATATCTG	1620	
QY	3258	TG	AGTTCTGAACGCTACTGAGATGACCTGTGACGCGCCGCTCTCGGTGTCTTGA	3317	
Db	1621	TG	AGTTCTGAACGCTACTGAGATGACCTGTGACGCGCCGCTCTCGGTGTCTTGA	1680	
QY	3318	CC	ACAGATCAAGCTGACCGAGAGGCCCGAGAGATTTGGCTTCATCTCTGACAACGTCTCA	3377	
Db	1681	CC	ACAGATCAAGCTGACCGAGAGGCCCGAGAGATTTGGCTTCATCTCTGACAACGTCTCA	1740	
QY	3378	GT	CCCTGCTCATCTCAACAGACCACTTCACTTATCTCCAACCCGGTGTTTGAGC	3437	
Db	1741	GT	CCCTGCTCATCTCAACAGACCACTTCACTTATCTCCAACCCGGTGTTTGAGC	1800	
QY	3438	CT	TTGATCCCTCAGGAATCTTGAAGCTCAAGGCTGGACAGCCCATCATCTCAAAAGGCGAA	3497	
Db	1801	CT	TTGATCCCTCAGGAATCTTGAAGCTCAAGGCTGGACAGCCCATCATCTCAAAAGGCGAA	1860	
QY	3498	GA	ACCTGATCCCGCTGTGGCTGGGGGCAACTGAAGCTGAACCTACCTGTGTGCTGG	3557	
Db	1861	GA	ACCTGATCCCGCTGTGGCTGGGGGCAACTGAAGCTGAACCTACCTGTGTGCTGG	1920	
QY	3558	GGA	GAGACCGGTGACCGTGTACAGTGTCCAGGTGACTCTCGAGTCCGCCCACT	3617	
Db	1921	GGA	GAGACCGGTGACCGTGTGTACAGTGTCCAGGTGACTCTCGAGTCCGCCCACT	1980	
QY	3618	CAT	CGGACAGGCAAAAGTATGGCCCGTGTGCTGTGGCATGAGATCTCCCGGGAGATGT	3677	
Db	1981	CAT	CGGACAGGCAAAAGTATGGCCCGTGTGCTGTGGCATGAGATCTCCCGGGAGATGT	2040	
QY	3678	GT	ACATTTCCCGGACAGCCCGCTCAAGCTTCGCCGCACTGTACAGATGCAGTGGCTGG	3737	

QY	Db	RESULT 4	LOCUS	DEFINITION	ACCESSION	VERSION	KEYWORDS	ORGANISM	REFERENCE	TITLE	JOURNAL	MEDLINE	PUBMED	REFERENCE	AUTHORS	TITLE	JOURNAL	REMARK	COMMENT
2041		GTGACATTGCCCCCGAGACGCCCCGTCACGCCCTGCCCGCATGTCAGCATCCAGTCCAGTCCG	2160																
3738		CGGCGCCCTCATCATATTTTCATGTCGTGGCCGCGTCTCATTTGCTCTATAAAGCAAGTCCCGCGA	3797																
2101		CGGCTCTCTCATATTTTTCATGTCGTGGCCGCGTCTCATTTGCTCTATAAAGCAAGTCCCGCGA	2160																
3798		AACTGACCTTCACGCTGGAAGCGGCTGCAGATGCAGATGGAACAACCTGGAAGTCCGCTGTCG	3857																
2161		AAATGACCTTCACGCTGGAAGCGGCTGCAGATGCAGATGGAACAACCTGGAAGTCCGCTGTCG	2220																
3858		CGTGAAGTGCAGAGAG 3874																	
2221		CTTGAGATGCAGAGAG 2237																	
BC028744		2015 bp	mRNA	linear	PRI 08-OCT-2003														
BC028744		Homo sapiens hypothetical protein DKFZP434G0625, mRNA (cDNA clone																	
BC028744		MGC:34033 IMAGE:4828816), complete cds.																	
BC028744.2		GI:34190713																	
MGC.																			
Homo sapiens (human)																			
Homo sapiens																			
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;																			
Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.																			
1 (bases 1 to 2015)																			
Straussberg, R.L., Feingold, E.A., Grouse, L.H., Derge, J.G.,																			
Klausner, R.D., Collins, F.S., Wenzel, L., Sherman, C.M., Shat, N.K.,																			
Altschul, S.F., Zeeberg, B., Bluetow, K.H., Schaefer, C.F., Bhat, N.K.,																			
Hopkins, R.F., Jordan, H., Moore, T., Max, S.T., Wang, J., Heien, F.,																			
Diatchenko, L., Marusina, K., Farmer, A.A., Rubin, G.M., Hong, L.,																			
Scapleton, M., Soares, M.B., Bonaldo, M.F., Casavant, T.L.,																			
Schoezy, T.E., Brownstein, M.J., Usdin, T.B., Toshiyuki, S.,																			
Carninci, P., Prange, C., Raha, S.S., Loquellano, N.A., Peters, G.J.,																			
Abrahamson, R.D., Mulhany, S.J., Bosak, S.A., McEwen, P.J.,																			
McKernan, K.D., Malek, J.A., Gamarallu, P.H., Richards, S.,																			
Worley, K.C., Hale, S., Garcia, A.M., Gay, L.D., Huyls, S.M.,																			
Vallalun, D.K., Muzny, D.M., Sodergren, E.J., Lu, X., Gibbs, R.A.,																			
Faley, J., Helton, E., Kettelman, M., Madan,																			

through the I.M.A.G.E. Consortium/LLNL at: <http://image.llnl.gov>
 Series: IRAC Plate: 47 Row: D Column: 11.
 Location/Qualifiers
 1. 2015
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /clone="MGC:34033 IMAGE:4828616"
 /tissue_type="Testis"
 /clone_id="NH1_MGC_97"
 /lab_host="DH10B"
 /note="Vector: pBluescript"

gene
 1. 2015
 /gene="DKFZp434G0625"
 /db_xref="LocusID:91584"
 378. 1946
 /product="hypothetical protein DKFZp434G0625"
 /protein_id="AAH28744.1"
 /db_xref="GI:22382105"
 /db_xref="LocusID:91584"
 /translation="MKAMPNWTCLSHLMVWGSSLTITROPAPLSKORSFYTER
 GEPAEGFNHLYVDETFHLYLGAVNRIYKLSYDLYVTHETSPEDNEKCPRIYV
 TCNEPLITNNVNMILIDYKENRLIACGSLYOGICLRLBEDLFLGSPYRKEHYL
 SGVNSGSPVGVIVSYNLDKFLIATVDGKPEYPTLSRKLTKNSPADDMFAVY
 HDEFVAMIKIPSDIPTIIPDPIYVYGFSSGNFYFLIOPEWVSPGSTKEQVY
 TSKLVRLKEDTAFNSYVEYVIGICERSGVYRLLQAAVLSKAVAYRGLGAVDDI
 LPTVSKGQRKMKSLDESLCTFIKIQINDRIKELSGYEGTGLDAMVXVDIP
 CSSELLITDINFGCDLDMNAPLGVSDMVRGIPVTERPDMDTVIAYVWNLAFPGT
 KSGKLKSPFGFPGGGLTQEMTGVAGDPFPAANIASQPMCLVTLQSSSKAISDQVQ
 PLLCFLNVPSNS"
 528. 1793
 /note="Sema; Region: semaphorin domain"
 /db_xref="CDD:smart00630"

misc_feature

528. 1793
 /note="Sema; Region: semaphorin domain"
 /db_xref="CDD:smart00630"

ORIGIN

Query Match 24.1%; Score 1371; DB 9; Length 2015;
 Best Local Similarity 100.0%; Pred. No. 0;
 Matches 1371; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

1 ATGAAGGCATGCCTGGAATGACCTGCTCTCTCCACCTCTCATGTGAGCATG 60
 378 ATGAAGGCATGCCTGGAATGACCTGCTCTCTCCACCTCTCATGTGAGCATG 437
 61 GGCTCTCTCACTTTGCTCAACCCGCGACGACCCCGCTGTCCGACGAGCGGTCATT 120
 438 GGCTCTCTCACTTTGCTCAACCCGCGACGACCCCGCTGTCCGACGAGCGGTCATT 497
 121 GTCCATTCGAGGAGGAGCCGCGAGGGTTTCATCACTGCTGTGATGAGAGACA 180
 498 GTCCATTCGAGGAGGAGCCGCGAGGGTTTCATCACTGCTGTGATGAGAGACA 557
 181 GACACATTTACTTGGGGGCGGTCAATGCAATTTTCAAGCTCTCCAGCGACTGAAGTCT 240
 558 GACACATTTACTTGGGGGCGGTCAATGCAATTTTCAAGCTCTCCAGCGACTGAAGTCT 617
 241 TTGTGAGCGCATGAGACGAGGCGGAGCGAGGACCAACCCCAAGTTTACCAACCCCGCATC 300
 618 TTGTGAGCGCATGAGACGAGGCGGAGCGAGGACCAACCCCAAGTTTACCAACCCCGCATC 677
 301 GTCCAGACCTGCATGAGCCCTGACACCAACCAATGTCAACCAAGATGCTCTCAT 360
 678 GTCCAGACCTGCATGAGCCCTGACACCAACCAATGTCAACCAAGATGCTCTCAT 737
 361 GACTCAAGAGGAGACAGGCTGATTTCCCTGTGGAGCGGTGACCAAGCATGTGCAAGCTG 420
 738 GACTCAAGAGGAGACAGGCTGATTTCCCTGTGGAGCGGTGACCAAGCATGTGCAAGCTG 797
 421 CTGAGGCTGAGGAGCCTCTTCAAGCTGTGGAGAGCCTTATCATTAAGAAAGAGCACTATCTG 480
 798 CTGAGGCTGAGGAGCCTCTTCAAGCTGTGGAGAGCCTTATCATTAAGAAAGAGCACTATCTG 857
 481 TAGGCTGACGAGAGGCGCTCACTTTTGAAGTGAATCGTCTCTACAGCAACCTGGAT 540

858 TGAAGTGTCAACGAGAGCGGCTCACTGTTTGGATGATCGTCTCTCAACCAACCTGGAT 917
 541 GACAAGCTGTTCATTCGACGAGCGAGTGAATGGGAAGCCCGAGTATTTTCCACCATCTCC 600
 918 GACAAGCTGTTCATTCGACGAGCGAGTGAATGGGAAGCCCGAGTATTTTCCACCATCTCC 977
 601 AGCCGGAACCTGACCAAGAACTGAGAGCGGATGAGTGTGGGTACGTCTTCATGAT 660
 978 AGCCGGAACCTGACCAAGAACTGAGAGCGGATGAGTGTGGGTACGTCTTCATGAT 1037
 661 GAGTTCGTGGCTGATGATTAAGATCCCTTGGACACTTACCATCATCTCTGACTTT 720
 1038 GAGTTCGTGGCTGATGATTAAGATCCCTTGGACACTTACCATCATCTCTGACTTT 1097
 721 GATATCTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 780
 1098 GATATCTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1157
 781 CCTGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 840
 1158 CCGTGAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1217
 841 GTGAGGCTTTGACAGAGAGACACAGCTTCACTCCATGATGATGATGATGATGATGATGATGAT 900
 1218 GTGAGGCTTTGACAGAGAGACACAGCTTCACTCCATGATGATGATGATGATGATGATGATGAT 1277
 901 GAGCGCAGTGGGATGAGATACCGGCTGCTGACGCTGCTTCACTCTGTCACCAAGCGGAGGCC 960
 1278 GAGCGCAGTGGGATGAGATACCGGCTGCTGACGCTGCTTCACTCTGTCACCAAGCGGAGGCC 1337
 961 GTGCTTGGCAGAGACCTTTGAGAGTCCATTCAGATGATGATGATGATGATGATGATGATGATGAT 1020
 1338 GTGCTTGGCAGAGACCTTTGAGAGTCCATTCAGATGATGATGATGATGATGATGATGATGATGAT 1397
 1021 AAGGCGCAGAAAGCGGAAATGAAATCCCTGATGATGATGATGATGATGATGATGATGATGATGAT 1080
 1398 AAGGCGCAGAAAGCGGAAATGAAATCCCTGATGATGATGATGATGATGATGATGATGATGATGAT 1457
 1081 AAGCAGATTAATGACCGGATTAAGAGCGGCTGACGCTTGTACCGGAGGAGAGGACAG 1140
 1458 AAGCAGATTAATGACCGGATTAAGAGCGGCTGACGCTTGTACCGGAGGAGAGGACAG 1517
 1141 CTGACCTGGGCTGGCTCAAGGTGAGAGACATCCCTGACAGTGGCTCTTAACATT 1200
 1518 CTGACCTGGGCTGGCTCAAGGTGAGAGACATCCCTGACAGTGGCTCTTAACATT 1577
 1201 GACGATTAATCTTGTGGCTTGGCTGACATGATGATGATGATGATGATGATGATGATGATGATGAT 1260
 1578 GACGATTAATCTTGTGGCTTGGCTGACATGATGATGATGATGATGATGATGATGATGATGATGAT 1637
 1261 GGAATTTCCGCTTACGAG 1320
 1638 GGAATTTCCGCTTACGAG 1697
 1321 AAGAACCACTCTCTGAGCTTTGTGGGACCAAAAGTGGCAAGCTGAAGAG 1371
 1698 AAGAACCACTCTCTGAGCTTTGTGGGACCAAAAGTGGCAAGCTGAAGAG 1748

RESULT 5
 AK123428 3310 bp mRNA linear PRI 09-SEP-2003
 LOCUS AK123428
 DEFINITION Homo sapiens CDNA FLJ41434 f1s, clone BHHF2007616, moderately
 similar to Plectin 2.
 ACCESSION AK123428
 VERSION AK123428.1 GI:34528974
 KEYWORDS oligo capping; f1s (full insert sequence).
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.

REFERENCE 1

AUTHORS

Kawakami, B., Sugiyama, A., Takemoto, M., Sugiyama, T., Irie, R., Otsuki, T., Sato, H., Wakamatsu, A., Ishii, S., Yamamoto, J., Isono, Y., Kawai, H., Y., Saito, K., Nishikawa, T., Kimura, K., Yamashita, H., Matsuo, K., Nakamura, Y., Sekine, M., Kikuchi, H., Kanda, K., Ogasuma, M., Murakawa, K., Kanehori, K., Takahashi, Fujii, A., Ogihara, A., Suzuki, Y., Sugano, S., Nagahari, K., Matsuo, Y., Nagai, K., and Isoe, T.

TITLE

NEDO human cDNA sequencing project

JOURNAL

Unpublished

REFERENCE

2 (bases 1 to 3310)

AUTHORS

Isoe, T. and Yamamoto, J.

JOURNAL

Submitted (15-JUL-2003) Takao Isoe, FUJ Project (HRI Team); 2-6-7

COMMENT

Kazusa-Kamatari, Kisarazu, Chiba 292-0818, Japan
(E-mail: genomics@kizai.co.jp, Tel: 81-438-52-3975, Fax: 81-438-52-3986)
NEDO human cDNA sequencing project supported by Ministry of Economy, Trade and Industry of Japan; cDNA full insert sequencing: Research Association for Biotechnology (RAB); cDNA library construction: Helix Research Institute (HRI) (supported by Japan Key Technology Center etc.); 5'- & 3'- end one pass sequencing: RAB, HRI, and Biotechnology Center, National Institute of Technology and Evaluation; clone selection for full insert sequencing: HRI and RAB; annotation: HRI and RAB.

FEATURES

Location/Qualifiers

1..3310
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="BRH1P207616"
/issue_type="hippocampus"
/clone_lib="BRH1P2"
/note="Cloning vector: pME18SFL3"
136..1614
/note="unlabeled protein product"
/codon_start=1
/protein_id="BAC85615.1"
/db_xref="gi:34528975"
/translation="MKAMPWMTCLSHLVMGMSSTLITRPAPI.SOKRSPVTFR
GEPAGEVHLVDERTGHIYGVANRI.YKLSKLVLTETGDEBNPACYPRLVQ
TGNPILTTNNVNNKMLIDYKRNLIAGSLYOGICLLEDELFIDGPEYHKKH
SGVNSGSEVGVIVSYNDUKLFIATVNDGKPEPTTSRLKTSNSADMGAVVF
HDEPVASIKIPSDFTIIPFDIYIVGSSGNFYFTLQPMNSPDSITKEOV
TSKLVLCKEIDTANSYVEVETGERSGVERLQAAVLSRGAIVLRTIGVFPDD
LFTVFSKQKRMKSLDESACIFILQINDRIKERQSCYRGGITLDMLVKQIP
CSAALLITDNGCLDMNAPLGVSDMVRGIPVETEDRMTSVIAYVYKXNSLAFVGT
KSGKLKVPGLSCLPTLELQTPRSHRATVTELLFSSCSSN"

CDS

ORIGIN

Query Match 24.1%; Score 1371, DB 9; Length 3310;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1371; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

1 ATGAAGCAGTCCCTGGAACCTGACCTGCTCTCCACCTCTCATGATGAGGATG 60
136 ATGAAGCAGTCCCTGGAACCTGACCTGCTCTCTCCACCTCTCATGATGAGGATG 195
61 GGCCTCTCACTTCTGCTCAACCGGCAAGCCCGGCTGTCCCGAAGCAGGCTATT 120
196 GGCCTCTCACTTCTGCTCAACCGGCAAGCCCGGCTGTCCCGAAGCAGGCTATT 255
121 GTCAATTCGAGAGAGAGCCGCGAGAGGATTTCAATCACTGTGTGATGAGAGCA 180
256 GTCAATTCGAGAGAGAGCCGCGAGAGGATTTCAATCACTGTGTGATGAGAGCA 315
181 GGAACATTTACTTGGGGGCGCTCAATCGATTCAAGCTCTCCAGGACCTGAAGTC 240
316 GGAACATTTACTTGGGGGCGCTCAATCGATTCAAGCTCTCCAGGACCTGAAGTC 375
241 TTGGGAGCGATGAGAGAGGCGCGAGAGCAAGCCCAAGTGTATCCAGCCGCGATC 300
376 TTGGGAGCGATGAGAGAGGCGCGAGAGCAAGCCCAAGTGTATCCAGCCGCGATC 435
301 GTCCAGACCTGCAATGAGCCCTGACCAACCAACAAATGTCAACAGATGCTCTCAT 360

436 GTCCAGACCTGCAATGAGCCCTGACCAACCAACAAATGTCAACAGATGCTCTCAT 495
361 GACTCAAGAGAGAGAGGCTGATGCTCTGTGGAGCCCTGTACCAAGGATGTGCAAGCTG 420
436 GACTCAAGAGAGAGAGGCTGATGCTCTGTGGAGCCCTGTACCAAGGATGTGCAAGCTG 555
421 CTGAGGCTGTGAGAGAGCTTTTCAAGCTGGGGAGACCTTTATCAATGAGAGACATATGTG 480
556 CTGAGGCTGTGAGAGAGCTTTTCAAGCTGGGGAGACCTTTATCAATGAGAGACATATGTG 615
481 TCAGGTGTCAACGAGAGGAGCTCAGTCTTTGAGATGATGCTCTCTCAAGCAACCTGTGAT 540
616 TCAGGTGTCAACGAGAGGAGCTCAGTCTTTGAGATGATGCTCTCTCAAGCAACCTGTGAT 675
541 GACAACTGTTTCAATGACAGGCGAGTGTATGGAGAGCCCGAGATATTTTCCACCATCTCC 600
676 GACAACTGTTTCAATGACAGGCGAGTGTATGGAGAGCCCGAGATATTTTCCACCATCTCC 735
601 AGCCGGAACCTGACCAAGAGCTCTGAGGCGAGTGTGATGCTCTCTCAAGCAACCTGTGAT 660
736 AGCCGGAACCTGACCAAGAGCTCTGAGGCGAGTGTGATGCTCTCTCAAGCAACCTGTGAT 795
661 GAGTTCGTGCTGTGATGATTAAGATCCCTTGTGAGACCTTCAACCTCATCTCTGATCTT 720
796 GAGTTCGTGCTGTGATGATTAAGATCCCTTGTGAGACCTTCAACCTCATCTCTGATCTT 855
721 GATATCTCTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 780
856 GATATCTCTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 915
781 CCTGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 840
916 CCTGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 975
841 GTGAGGCTTTGAG 900
976 GTGAGGCTTTGAG 1035
901 GAGCGAGTGGGATGAGATCCGCTGTGACAGGCTGTGACAGGCTGTGACAGGCTGTGAC 960
1036 GAGCGAGTGGGATGAGATCCGCTGTGACAGGCTGTGACAGGCTGTGACAGGCTGTGAC 1095
961 GTGCTTGGAG 1020
1096 GTGCTTGGAG 1155
1021 AAGGGCCAG 1080
1156 AAGGGCCAG 1215
1081 AAGCAGATTAATGACCGCATTTAAGAGAGGCGCTGAGTCTGTGATCCGGGGCGAGGGACG 1140
1216 AAGCAGATTAATGACCGCATTTAAGAGAGGCGCTGAGTCTGTGATCCGGGGCGAGGGACG 1275
1141 CTGAGCTTGGCGCTGAGTCAAGGTGAAGAGATCCCTGAGAGAGAGTGGCTCTTAACAT 1200
1276 CTGAGCTTGGCGCTGAGTCAAGGTGAAGAGATCCCTGAGAGAGAGTGGCTCTTAACAT 1335
1201 GACGATTAATGATGAGCTGAGTCAAGGTGAAGAGATCCCTGAGAGAGAGTGGCTCTTAAC 1260
1336 GACGATTAATGATGAGCTGAGTCAAGGTGAAGAGATCCCTGAGAGAGAGTGGCTCTTAAC 1395
1261 GGAATTCGCGCTTCAAG 1320
1396 GGAATTCGCGCTTCAAG 1455
1321 AAGAACCACTGCTGTGCTTTGTGGGACCAAAAGTGGCAAGCTGAAGAG 1371
1456 AAGAACCACTGCTGTGCTTTGTGGGACCAAAAGTGGCAAGCTGAAGAG 1506

AC009785/c 179145 bp DNA linear PRI 26-MAY-2002
 LOCUS AC009785
 DEFINITION Homo sapiens chromosome 7 clone RP11-198E23, complete sequence.
 ACCESSION AC009785
 VERSION AC009785.8 GI:21217401
 KEYWORDS HTG.
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.
 REFERENCE AUTHORS
 1 (bases 1 to 179145)
 Kaul, R.K., Olson, M.V., Zhou, Y., James, R.A., Rouse, G., Wu, Z.,
 Saenphimmachak, C., Phelps, K.A., Buckley, D., Kibukawa, M., Raymond, C.
 and Haugen, E.D.
 TITLE Direct Submission
 JOURNAL Unpublished
 REFERENCE AUTHORS
 2 (bases 1 to 179145)
 Bubb, K.L., Desmarais, C.L., Ramsey, S.A. and Hudley, R.M.
 TITLE Direct Submission
 JOURNAL Submitted (01-SEP-1999) Human Genome Center, University of
 Washington, Box 352145, Seattle, WA 98195, USA
 3 (bases 1 to 179145)
 Kaul, R.K., Olson, M.V., Zhou, Y., James, R.A., Raymond, C.,
 Cleland, D., Ivey, R.G. and Haugen, E.D.
 TITLE Direct Submission
 JOURNAL Submitted (20-JUN-2001) Genome Center, University of Washington,
 Box 352145, Seattle, WA 98195, USA
 4 (bases 1 to 179145)
 Kaul, R.K., Olson, M.V., Zhou, Y., James, R.A., Rouse, G., Wu, Z.,
 Saenphimmachak, C., Phelps, K.A., Raymond, C. and Haugen, E.D.
 TITLE Direct Submission
 JOURNAL Submitted (06-APR-2002) Genome Center, University of Washington,
 Box 352145, Seattle, WA 98195, USA
 5 (bases 1 to 179145)
 Kaul, R.K., Olson, M.V., Zhou, Y., James, R.A., Rouse, G., Wu, Z.,
 Saenphimmachak, C., Phelps, K.A., Buckley, D., Kibukawa, M., Raymond, C.
 and Haugen, E.D.
 TITLE Direct Submission
 JOURNAL Submitted (26-MAY-2002) Genome Center, University of Washington,
 Box 352145, Seattle, WA 98195, USA
 On May 26, 2002 this sequence version replaced gi:20066281.
 COMMENT
 Center: University of Washington Genome Center
 Center Code: UWGC
 Web site: <http://www.genome.washington.edu>
 Contact: uwgchgs@u.washington.edu
 Project Information
 Center project name: chr-7
 Center clone name: RP11-198E23 (djs380)
 Summary Statistics
 Sequencing vector: plasmid; X52328, 100% of reads
 Chemistry: Dye-terminator Big Dye; 100% of reads
 Assembly program: Phrap; version 0.990319
 Consensus quality: 178850 bases at least Q40
 Consensus quality: 179053 bases at least Q30
 Consensus quality: 179141 bases at least Q20
 Insert size: 179145; sum-of-contigs
 Quality coverage: 11.8x in Q20 bases; sum-of-contigs

Overlapping Sequences:
 5' : RP11-341C17 (UWGC:djs301) AC011625 73322-bp overlap
 3' : RP11-335B7 (UWGC:djs42) AC026239 38206-bp overlap

Sequence Quality Assessment:
 This entry has been annotated with sequence quality
 estimates computed by the Phrap assembly program.
 All manually edited bases have been reduced to quality zero.
 Quality levels above 40 are expected to have less than
 1 error in 10,000 bp.
 Base-by-base quality values are not generally visible from the
 GenBank flat file format but are available as part
 of this entry's ASN.1 file.

This sequence was finished as follows unless otherwise noted:
 all regions were either double-stranded or sequenced with an
 alternate chemistry or covered by high quality data (i.e., Phred
 quality >= 30); an attempt was made to resolve all sequencing
 problems, such as compressions and repeats; all regions were
 covered by at least one plasmid subclone or more than one M3
 subclone; and the assembly was confirmed by restriction digest.

Sequence Validation:

This sequence has been validated by Multiple Complete Digest
 fingerprinting. Comparison of the experimentally derived digest
 fragments with sequence-predicted fragments is given below.
 The electronically-digested sequence consists of both insert and
 vector, in order to accurately represent the entire circular BAC.
 Small fragments below a variable cutoff (approximately 400-800 bp)
 are not resolved in the fingerprint and hence do not appear
 in the table. There are no significant remaining discrepancies
 between the experimental and predicted values. Uniquely ordered
 fragments are separated by dashed lines.

EcoRI				HindIII				BglII			
SeqDerMap	FngPrnt	SeqDerMap	FngPrnt	SeqDerMap	FngPrnt	SeqDerMap	FngPrnt	SeqDerMap	FngPrnt	SeqDerMap	FngPrnt
8696	8850	399	<800	2304	2280						
6	<800	6382	6347	2067	2069						
1907	1895	512	<800	5518	5637						
1854	1895	449	<800	6918	7002						
2849	2922	2322	2276	2029	2069						
3722	3806	5951	5995	2378	2425						
658	<800	7224	7291	2089	2069						
26398	26593	1005	1033	550	<800						
537	<800	1020	1033	4191	4187						
1542	1483	3445	3448	3839	3773						
1103	1096	2131	2126	1647	1646						
1928	2045	2237	2276	629	<800						
14029	13827	6625	6652	1932	1918						
6707	6816	6349	6347	18266	18041						
1884	1895	6669	6652	3099	3165						
6048	6029	1943	1916	2900	2928						
2150	2209	6376	6347	111	<800						
13944	13827	5398	5641	2425	2425						
4275	4287	599	<800	3767	3773						
2057	2209	4055	4127	1506	1485						
2293	2209	299	<800	3639	3773						
6448	6393	20065	20079	5635	5637						
9916	9943	6964	6903	6894	7002						
7673	7645	5766	5641	11819	11807						

115	<800	4281	4127	837	821
1490	1483	7844	7901	2058	2069
1071	1096	1032	1033	3280	3312
18341	18612	1591	1546	3780	3773
7618	7645	955	954	623	<800
1733	1700	330	<800	202	<800
1478	1483	1927	1916	2950	2928
7616	7645	6261	6347	10775	10621
561	<800	445	<800	12114	11807
19080	18612	3869	3877	9445	9461
114	<800	4148	4127	1694	1646
7	<800	3828	3877	3745	3773
		4001	4127	674	<800
		1561	1546	9893	9817
		3188	3192	3225	3165
		4960	4943	998	1011
		9109	8931	1318	1283
		929	954	94	<800
		7844	7901	796	821
		535	<800	111	<800
		5397	5408	1029	1011
		1905	1916	1491	1485
		113	<800	20564	20918
		5596	5641		
		750	735		
		1264	1248		

FEATURES

Location/Qualifiers

1..179145

/organism="Homo sapiens"

Query Match 20.9%; Score 1188; DB 9; Length 179145;

Best Local Similarity 100.0%; Pred. No. 0;

Matches 1188; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db	61616	GGACGCAATTACTTGGGGGCGGTCAATCGGATTTACAAGCTCTCCAGGACCTGAGAGTC	61557
Qy	241	TTGGTGAAGCATGAGACAGGGCCGGACGAGACAAACCCCAAGTGTATCCACCCCGCATC	300
Db	61556	TTGGTGAAGCATGAGACAGGGCCGGACGAGACAAACCCCAAGTGTATCCACCCCGCATC	61497
Qy	301	GTCCAGACCTGCAATGAGCCCTTGAACACACCAACATGTCAACAATGCTCTCTATA	360
Db	61496	GTCCAGACCTGCAATGAGCCCTTGAACACACCAACATGTCAACAATGCTCTCTATA	61437
Qy	361	GACTTCAAGAGAAACAGGCTGATTTGCTGTGGAGCTGTACCAAGCATTTGCAAGCTG	420
Db	61436	GACTTCAAGAGAAACAGGCTGATTTGCTGTGGAGCTGTACCAAGCATTTGCAAGCTG	61377
Qy	421	CTGAGGCTGAGAGACCTTTCAAGCTGGGGAGGCTTATCATTAAGAGAGACATATCTG	480
Db	61376	CTGAGGCTGAGAGACCTTTCAAGCTGGGGAGGCTTATCATTAAGAGAGACATATCTG	61317
Qy	481	TCAGGTGTCACGAGAGCGCTCAGTCTTTGAGTATCGTCTCCACAGCACTTGAT	540
Db	61316	TCAGGTGTCACGAGAGCGCTCAGTCTTTGAGTATCGTCTCCACAGCACTTGAT	61257
Qy	541	GACAGCTGTTCAATTGCCAGGCGCATGTGATGGAGAGCCGAGTATTTTCCACCATCTCC	600
Db	61256	GACAGCTGTTCAATTGCCAGGCGCATGTGATGGAGAGCCGAGTATTTTCCACCATCTCC	61197
Qy	601	AGCCGGAACCTGACCAAGAACTCTGAGCGGATGCGATGTTGCGGTACGTCTTCATGAT	660
Db	61196	AGCCGGAACCTGACCAAGAACTCTGAGCGGATGCGATGTTGCGGTACGTCTTCATGAT	61137
Qy	661	GAGTTCGTGCTGATGATGATTAAGTCCCTTCGAGACCTTCACCATATCCCTGACTTC	720
Db	61136	GAGTTCGTGCTGATGATGATTAAGTCCCTTCGAGACCTTCACCATATCCCTGACTTC	61077
Qy	721	GATATCTACTATGTCATGATGTTTGAAGTGGCACTTGTCTATCTTTTACCCTCCAA	780
Db	61076	GATATCTACTATGTCATGATGTTTGAAGTGGCACTTGTCTATCTTTTACCCTCCAA	61017
Qy	781	CCTGAGATGATGTCCTCCACAGGCTCCACCAAGAGAGAGTATATCATCCAGCTC	840
Db	61016	CCTGAGATGATGTCCTCCACAGGCTCCACCAAGAGAGAGTATATCATCCAGCTC	60957
Qy	841	GTGAGGCTTTGCAAGAGAGACAGGCTTCAACTCTATGTAGAGTGGCCCATTTGCTGT	900
Db	60956	GTGAGGCTTTGCAAGAGAGACAGGCTTCAACTCTATGTAGAGTGGCCCATTTGCTGT	60897
Qy	901	GAGCGCAGTGGGTGAGTACCGCTGCTGCAAGGCTGCTTCTCTCAAGCGGGGCC	960
Db	60896	GAGCGCAGTGGGTGAGTACCGCTGCTGCAAGGCTGCTTCTCTCAAGCGGGGCC	60837
Qy	961	GTGCTTGCAAGAGACCTTGGAGTCCATCCAGATGATGACCTGCTTCAACCGCTTCTCC	1020
Db	60836	GTGCTTGCAAGAGACCTTGGAGTCCATCCAGATGATGACCTGCTTCAACCGCTTCTCC	60777
Qy	1021	AGGCGCCAGAGCGGAAAAATGAATCCTGATGATGAGTGGCCCTGTGATCTTCACTTG	1080
Db	60776	AGGCGCCAGAGCGGAAAAATGAATCCTGATGATGAGTGGCCCTGTGATCTTCACTTG	60717
Qy	1081	AAGCAGATTAATGACCGATTAAGGAGGCGCTGAGCTTGTGTACCGGGGCGAGGGACG	1140
Db	60716	AAGCAGATTAATGACCGATTAAGGAGGCGCTGAGCTTGTGTACCGGGGCGAGGGACG	60657
Qy	1141	CTGACCTGGCCTGCTCAAGGTGAAGACATCCCTTCAGACAGTGC	1188
Db	60656	CTGACCTGGCCTGCTCAAGGTGAAGACATCCCTTCAGACAGTGC	60609

RESULT 7

AX833678

LOCUS

DEFINITION

Sequence 802 from Patent EP1347046.

AX833678

ACCESSION

VERSION AX833678.1 GI:39919813
 KEYWORDS
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 REFERENCE
 1 Isogai, T., Sugiyama, T., Otsuki, T., Makatsugu, A., Sato, H., Ishii, S., Yamamoto, J., I., Isono, Y., Hio, Y., Otsuka, K., Nagai, K., Irie, R., Tamechika, I., Seki, N., Yoshikawa, T., Otsuka, M., Nagahara, K. and Masuh, Y.
 TITLE Full-length cDNA sequences
 JOURNAL Patent: EP 1347046-A 802 24-SEP-2003;
 Research Association for Biotechnology (JP)
 FEATURES
 source 1..3666
 /organism="Homo sapiens"
 /mol_type="unassigned DNA"
 /db_xref="taxon:9606"
 ORIGIN
 Query Match 19.5%; Score 1108; DB 6; Length 3666;
 Best Local Similarity 99.9%; Pred. No. 0;
 Matches 1158; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
 QY 4527 TGCCACAGCCCGAGGAGTCCAGTAAGATCTCACTGACACCATCACTGAGTCA 4586
 1 TGCCACAGCCCGAGGAGTCCAGTAAGATCTCACTGACACCATCACTGAGTCA 60
 Db 4587 GAGAAAGATTTGGATGCTTCAAGAAATGTCCTGCTCCACCGGCGCAAGCTGC 4646
 61 GAGAAAGATTTGGATGCTTCAAGAAATGTCCTGCTCCACCGGCGCAAGCTGC 120
 QY 4647 AGATATGATCTGAGTGGCGACAGAGAGTGGGCGAGAGTATCTTGACAGATCA 4706
 121 AGATATGATCTGAGTGGCGACAGAGAGTGGGCGAGAGTATCTTGACAGATCA 180
 Db 4707 CATACACCAAGATTGAGAAATGATTGGAGCCATGTAACAACCTGCGCCATCAAGAT 4766
 181 CATACACCAAGATTGAGAAATGATTGGAGCCATGTAACAACCTGCGCCATCAAGAT 240
 QY 4767 GCCAGATGTTCCCTGGTGGCATTAAGTGTCAAGCAGAGTCAAGCCTATACCACTGAA 4826
 241 GCCAGATGTTCCCTGGTGGCATTAAGTGTCAAGCAGAGTCAAGCCTATACCACTGAA 300
 Db 4827 CAATCTCAAGCTCTCCAGAGCTTCAAGCAAGTAAATATGAAACATGATCCGGTACACGG 4886
 301 CAATCTCAAGCTCTCCAGAGCTTCAAGCAAGTAAATATGAAACATGATCCGGTACACGG 360
 QY 4887 CAGCCCGGAGAGCTCCGCTCAGAGACACCTATGATCACTCTGACCTGAGAGTGAAT 4946
 361 CAGCCCGGAGAGCTCCGCTCAGAGACACCTATGATCACTCTGACCTGAGAGTGAAT 420
 Db 4947 CAAGATGTGACCTTAAGTAAAGAACACAGAGACCGAGACCGAAGAGAGGGGGACGGGG 5006
 421 CAAGATGTGACCTTAAGTAAAGAACACAGAGACCGAGACCGAAGAGAGGGGGACGGGG 480
 QY 5007 GAGCAAGATGTTGTTGAATCTACCTGACCGGACTCCTGAGCCACTAAGGGGACACTGCA 5066
 481 GAGCAAGATGTTGTTGAATCTACCTGACCGGACTCCTGAGCCACTAAGGGGACACTGCA 540
 Db 5067 GAAATTTGTGATGATCTTTGAGACCATTTGAGACCGGACACCGGCTCTGCGCT 5126
 541 GAAATTTGTGATGATCTTTGAGACCATTTGAGACCGGACACCGGCTCTGCGCT 600
 QY 5127 GCCCTGGGCGCATCAAGTACATGTTTGAATCTTCTGATGAGAGGCTGATTAACATGAGCAT 5186
 601 GCCCTGGGCGCATCAAGTACATGTTTGAATCTTCTGATGAGAGGCTGATTAACATGAGCAT 660
 Db 5187 TCATGACCCGACGTCGCGCATACCTGGAAGAGCAATTCCTGCGCCCTGAGAGTTTGGGT 5246
 661 TCATGACCCGACGTCGCGCATACCTGGAAGAGCAATTCCTGCGCCCTGAGAGTTTGGGT 720

QY 5247 CAACATGATCAGAACCCGACAGTTTGTGTTGACATCCATAGAACAGATCAGACAGC 5306
 Db 721 CAACATGATCAGAACCCGACAGTTTGTGTTGACATCCATAGAACAGATCAGACAGC 780
 QY 5307 CTGCTCTCTGTGTGTGCTGACATCTTCAATGAGTCTTCTCTCCAGCTCAGACAGCGGCT 5366
 Db 781 CTGCTCTCTGTGTGTGCTGACATCTTCAATGAGTCTTCTCTCCAGCTCAGACAGCGGCT 840
 QY 5367 GGGCAGAGACTCGCCCTCCAAAGAGTGTGTATGCGCAGAGACATCCCGAGTACAGAA 5426
 841 GGGCAGAGACTCGCCCTCCAAAGAGTGTGTATGCGCAGAGACATCCCGAGTACAGAA 900
 QY 5427 TTGGGTGAGAGGATTAATCAACATGAGAGAGATCCAGCCATCAGCCACAGACAT 5486
 901 TTGGGTGAGAGGATTAATCAACATGAGAGAGATCCAGCCATCAGCCACAGACAT 960
 Db 5487 GAAAGCATACCTGCTGAGAGAGTCCGAGATGACATGATGATTAACACCATGAGTGC 5546
 961 GAAAGCATACCTGCTGAGAGAGTCCGAGATGACATGATGATTAACACCATGAGTGC 1020
 QY 5547 ACTCTCAGAGATCTTCTCTATATGTTGGCAATACAGCAGAGATCTTGGACCTTGA 5606
 1021 ACTCTCAGAGATCTTCTCTATATGTTGGCAATACAGCAGAGATCTTGGACCTTGA 1080
 QY 5607 CCAAGATGACCAAGTGTGGAGAGCAAGAACTGGCTCAAACTGAAACAGTCAATACCT 5666
 1081 CCAAGATGACCAAGTGTGGAGAGCAAGAACTGGCTCAAACTGAAACAGTCAATACCT 1140
 QY 5667 CATGAGCTTACAGAGCTGA 5685
 Db 1141 CATGAGCTTACAGAGCTGA 1159
 RESULT 8
 AK095606
 LOCUS
 DEFINITION Homo sapiens cDNA FLJ38287 f1s, clone FCBBF3008362, moderately similar to PLEXIN 4 PRECURSOR.
 AK095606
 ACCESSION AK095606.1 GI:21754899
 VERSION
 KEYWORDS Oligo capping; fls (full insert sequence).
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 REFERENCE
 1 Suzuki, O., Sasaki, N., Aotake, S., Shoji, T., Ichihara, T., Shiohara, N., Matsunoto, K., Hirano, M., Sano, S., Nemura, R., Yoshikawa, Y., Matsunura, Y., Moriya, S., Chiba, E., Momiyama, H., Onogawa, S., Kaeriyama, S., Satoh, N., Matsunawa, H., Takahashi, E., Katoka, R., Kuga, N., Kuroda, A., Satoh, I., Kamata, K., Takami, S., Terashima, Y., Watanabe, M., Sugiyama, T., Irie, R., Otsuki, T., Sato, H., Ota, T., Wakatsuki, A., Ishii, S., Yamamoto, J., Isono, Y., Kawai, Hio, Y., Saito, K., Nishikawa, T., Kimura, K., Yamashita, H., Matsuo, K., Nakamura, Y., Sekine, M., Kikuchi, H., Kanda, K., Waga, S., Murakawa, K., Kanehori, K., Takahashi, Fujii, A., Oshima, A., Sugiyama, A., Kawakami, B., Suzuki, Y., Sugano, S., Nagahara, K., Masuh, Y., Nagai, K. and Isogai, T.
 TITLE NED human cDNA sequencing project
 JOURNAL Unpublished
 REFERENCE 2 (bases 1 to 3666)
 AUTHORS Isogai, T. and Yamamoto, J.
 DIRECT SUBMISSION
 TITLE Submitted (04-JUN-2002) Takao Isogai, FLJ Project (HRI Team); 2-6-7 Kazusa-Kamatari, Kisarazu, Chiba 292-0812, Japan
 JOURNAL (E-mail:genomicshri.co.jp, Tel:81-438-52-3975, Fax:81-438-52-3986)
 COMMENT NED human cDNA sequencing project supported by Ministry of Economy, Trade and Industry of Japan; cDNA full insert sequencing; Research Association for Biotechnology (RAB); cDNA library construction; Helix Research Institute (HRI) (supported by Japan Key Technology Center etc.); 5'- & 3'-end one pass sequencing; RAB, HRI, and Biotechnology Center, National Institute of Technology and Evaluation; clone selection for full insert sequencing; HRI and

RAB; annotation: HRI and RAB.
Location/Qualifiers
source
1. 3666

/organism="Homo sapiens"
/mol_type="RNA"
/db_xref="taxon:9606"
/clone="FCB8F3008362"
/issue_type="brain"
/clone_id="FCB8F3"
/dev_stage="fetal"
/note="Cloning vector: pME18SFL3"
125..1159
/note="unnamed protein product"
/codon_start=1
/protein_id="BAC04587.1"
/db_xref="GI:21754900"
/translation="MDLEWRQSGGAMNLLDDEDTTKIENDMKRLNTLAHYOVPGSV
VALVSKQVTAYNAVNNSTVSRTSASRYENNTRTGSPDLSKSTPMTIPDLSSGYRW
HLVNHEDQDEQKEDRSGKMSVSHYLRLLATGTLQKQVDFEITFSTAHGSLP
LAIKYMPDFDEQKDHGIDPHVRHTWKNCLPLRFVWMIKNPQFVHDKNSLTD
ACLSVAVOTFMDSCSTSEHRLGKSPSNKLLVAKDIPSYKXWVERYSIDIGKMPAID
ODMAYVLAOSRHMNEFNMSALSEIFSYVGYKSEIILGPLDHDQCKOKLAYKLE
QVTLNLSLDS"

ORIGIN

Query Match 19.5%; Score 1108; DB 9; Length 3666;
Best Local Similarity 99.3%; Pred. No. 0;
Matches 1158; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 4527 TGGCAACAGCCCGGAGTCCCAAGTAAGATCTTCACTGTGACACCATCACTAGGTCAA 4586
DB 1 TGGCAACAGCCCGGAGTCCCAAGTAAGATCTTCACTGTGACACCATCACTAGGTCAA 60
QY 4587 GGAGAAATCTGTGATCCATCTTCAAGAAATGTGCTTCCACCGGCGGCAAGAGTGC 4646
DB 61 GGAGAAATCTGTGATCCATCTTCAAGAAATGTGCTTCCACCGGCGGCAAGAGTGC 120
QY 4647 AGATATGATGTGAGTGGCGACAAAGAAAGTGGGCGAAGATGATCTTTCAGAGTAAAG 4706
DB 121 AGATATGATGTGAGTGGCGACAAAGAAAGTGGGCGAAGATGATCTTTCAGAGTAAAG 180
QY 4707 CATCAACCAACCAAGATTGGAATGATGGAAGCACTGAACAACACTGAGCCACTACAGGT 4766
DB 181 CATCAACCAACCAAGATTGGAATGATGGAAGCACTGAACAACACTGAGCCACTACAGGT 240
QY 4767 GCCAGATGTTCCGTGTGGCATTAGTGTCCAGACAGGTGACAGCTTAAACGAGTAA 4826
DB 241 GCCAGATGTTCCGTGTGGCATTAGTGTCCAGACAGGTGACAGCTTAAACGAGTAA 300
QY 4827 CAATCCCACTCTCCGAGGACCTGAGCAAGTAATATGAAGAAATGATCCGGTAAACAGG 4886
DB 301 CAATCCCACTCTCTCCGAGGACCTGAGCAAGTAATATGAAGAAATGATCCGGTAAACAGG 360
QY 4887 CAGCCCGGACAGCTCCGCTCAGCGACACCTATGATCACTCTGAGAGTGAAGTGAAGT 4946
DB 361 CAGCCCGGACAGCTCCGCTCAGCGACACCTATGATCACTCTGAGAGTGAAGTGAAGT 420
QY 4947 CAAGATGTGGCACTTATGAAAGAACCAAGACAGGACCGAAGAGAGGAGGAGCCGGGG 5006
DB 421 CAAGATGTGGCACTTATGAAAGAACCAAGACAGGACCGAAGAGAGGAGGAGCCGGGG 480
QY 5007 GAGCAAGATGTGTCTGAATCTACTGACCCGACTCTGGGCACTTAAGGCAACACTGCA 5066
DB 481 GAGCAAGATGTGTCTGAATCTACTGACCCGACTCTGGGCACTTAAGGCAACACTGCA 540
QY 5067 GAAATTTTGTGATGACCTCTTTGAGACCATCTTACAGACCGGACACCGTGGCTCTGCCC 5126
DB 541 GAAATTTTGTGATGACCTCTTTGAGACCATCTTACAGACCGGACACCGTGGCTCTGCCC 600
QY 5127 GCCCTTGAGCAATCAAGTACATGTTGACTTCTGAGTACAGAGGCTGATAAACAATGACAT 5186
DB 601 GCCCTTGAGCAATCAAGTACATGTTGACTTCTGAGTACAGAGGCTGATAAACAATGACAT 660

QY 5187 TCATGACCCGACGCTCCGCCATACCTGGAAGAGCAATGTGCTGCTGAGGTTTGGGT 5246
DB 661 TCATGACCCGACGCTCCGCCATACCTGGAAGAGCAATGTGCTGCTGAGGTTTGGGT 720
QY 5247 CAACATGATCAAGAACCCGCGAGTTTGTGTTGATCATTCATTAAGAACAGCATCAAGACG 5306
DB 721 CAACATGATCAAGAACCCGCGAGTTTGTGTTGATCATTCATTAAGAACAGCATCAAGACG 780
QY 5307 CTGCTCTCTGT 5366
DB 781 CTGCTCTCTGT 840
QY 5367 GGGCAAGAGCTGCGCTCCCAACAGAGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 5426
DB 841 GGGCAAGAGCTGCGCTCCCAACAGAGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 900
QY 5427 TTGGGTGAGAGGATATTACTAGACATAGGAAAGTGCAGCCATCAAGCAAGACAT 5486
DB 901 TTGGGTGAGAGGATATTACTAGACATAGGAAAGTGCAGCCATCAAGCAAGACAT 960
QY 5487 GAACGATACCTGCTGTGAGCAGTCCCGGATGACATGAATGATTCAAACCATGAGTGC 5546
DB 961 GAACGATACCTGCTGTGAGCAGTCCCGGATGACATGAATGATTCAAACCATGAGTGC 1020
QY 5547 ACTTCAGAGATCTTCTCTTATGTGTGGCAATACAGGAGAGATCTCTTGAAGCTTGA 5606
DB 1021 ACTTCAGAGATCTTCTCTTATGTGTGGCAATACAGGAGAGATCTCTTGAAGCTTGA 1080
QY 5607 CCACGATGACAGTGTGTGGAAGAGAACTGCGCTCAACTAGAACAGTCAATACCT 5666
DB 1081 CCACGATGACAGTGTGTGGAAGAGAACTGCGCTCAACTAGAACAGTCAATACCT 1140
QY 5667 CATGAGCTTAAACAGCTGA 5685
DB 1141 CATGAGCTTAAACAGCTGA 1159

RESULT 9
AC011625/c 188090 bp DNA linear PRI 08-DEC-1999
LOCUS AC011625
DEFINITION Homo sapiens clone RP11-341C17 from 7q31, complete sequence.
ACCESSION AC011625
VERSION AC011625.2 GI:6539285
KEYWORDS HTG.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
1 (bases 1 to 188090)
Kaul, R.K., Yu, J., Wong, G.K.-S., Magness, C.L., Green, E.D., Green, P.
and Olson, M.V.
Large-scale MCD Mapping and Sequencing of Human Chromosome 7
Unpublished
2 (bases 1 to 188090)
Bubb, K.L., Desmarais, C.L., Ramsey, S.A. and Hubley, R.M.
Direct Submision
Submitted (08-OCT-1999) Human Genome Center, University of
Washington, Box 352145, Seattle, WA 98195, USA
3 (bases 1 to 188090)
Kaul, R.K. and Desmarais, C.L.
Direct Submission
Submitted (08-DEC-1999) Human Genome Center, University of
Washington, Box 352145, Seattle, WA 98195, USA
On Dec 8, 1999 this sequence version replaced gi:6016750.

----- Genome Center:
University of Washington Genome Center
Center code: UWGC
Web site: http://genome.washington.edu
Contact: uwgshelp@u.washington.edu
----- Project Information
Center project name: chr-7
Center clone name: djs301 (RP11-341C17)

Summary Statistics

Sequencing vector: M13, 100% of reads
Chemistry: Dye-Primer Bodipy, 93% of reads Chemistry:
Dye-terminator Big Dye, 7% of reads
Assembly program: Phrap, version 0.990319
Insert size: 188 094; sum-of-contigs
Quality coverage: 7.78x in Q20 bases; sum-of-contigs

Overlapping Sequences:

5 : UMGc:dj8302
3 : UMGc:dj8380

Sequence Quality Assessment:

This entry has been annotated with sequence quality estimates computed by the Phrap assembly program. All manually edited bases have been reduced to quality zero. Quality levels above 40 are expected to have less than 1 error in 10,000 bp. Base-by-base quality values are not generally visible from the GenBank flat file format but are available as part of this entry's ASN.1 file.

Double stranded (DS) coverage: 70.4%
DS or two chemistry coverage: 100.0%
Single stranded regions: 0

Sequence Validation:

This sequence has been validated by Multiple Complete Digest fingerprinting. Comparison of the experimentally derived digest fragments with sequence-predicted fragments is given below. The electronically-digested sequence consists of both insert and vector, in order to accurately represent the entire circular BAC. Small fragments below a variable cutoff (approximately 400-600 bp) are not resolved in the fingerprint and hence do not appear in the table. There are no significant remaining discrepancies between the experimental and predicted values. Uniquely ordered fragments are separated by dashed lines.

FP Seq FP Seq FP Seq
BglII EcoRI NsiI

2095.00	2061.00	8947.00	8687.00	729.00	732.00
12518.00	12347.00	844.00	812.00	1463.00	1480.00
5082.00	5007.00	4657.00	4593.00	5528.00	5438.00
2694.00	2672.00	3637.00	3613.00	3383.00	3371.00
5175.00	5083.00	1152.00	1147.00	1240.00	1220.00
1094.00	1093.00	11517.00	11411.00	4875.00	4844.00
2231.00	2189.00	16729.00	16537.00	4194.00	4094.00
16006.00	15888.00	13550.00	13383.00	2402.00	2353.00
6982.00	6668.00	5001.00	4900.00	1463.00	1439.00
8822.00	8629.00	6353.00	6107.00	2072.00	2056.00
11945.00	11811.00	3101.00	3027.00	5301.00	5143.00
10740.00	10587.00	3850.00	3719.00	1463.00	1408.00
2095.00	2062.00	1925.00	1935.00	729.00	706.00
7981.00	7820.00	6169.00	6045.00	10419.00	10239.00
2095.00	2101.00	6353.00	6206.00	4194.00	4107.00
1178.00	1141.00	1743.00	1714.00	1128.00	1118.00
3623.00	3525.00	2067.00	2052.00	622.00	630.00

621.00	581.00	559.00	569.00	1045.00	1052.00
6428.00	6341.00	544.00	510.00	7407.00	7133.00
6146.00	6023.00	2364.00	2354.00	3498.00	3389.00
769.00	736.00	660.00	633.00	4506.00	4428.00
4722.00	4640.00	1013.00	972.00	2402.00	2339.00
1178.00	1180.00	11962.00	11788.00	13094.00	12940.00
7083.00	6917.00	5201.00	5141.00	4378.00	4366.00
2095.00	2022.00	5201.00	5091.00	2025.00	1988.00
2429.00	2372.00	1925.00	1902.00	4037.00	3911.00
2095.00	2083.00	1925.00	1853.00	2402.00	2326.00
546.00	544.00	2967.00	2843.00	2072.00	2069.00
4240.00	4182.00	3850.00	3715.00	1609.00	1583.00
3913.00	3829.00	660.00	652.00	2922.00	2845.00
1656.00	1641.00	29666.00	26384.00	5301.00	5252.00
621.00	623.00	544.00	531.00	1747.00	1719.00
1940.00	1927.00	1551.00	1536.00	729.00	697.00
18696.00	18261.00	1152.00	1097.00	2646.00	2580.00
3153.00	3085.00	1925.00	1922.00	2402.00	2335.00
2966.00	2898.00	14113.00	14024.00	10419.00	10108.00
2429.00	2419.00	6876.00	6704.00	7407.00	7317.00
3913.00	3763.00	1925.00	1869.00	8785.00	8640.00
1524.00	1500.00	6169.00	6048.00	7052.00	6870.00
3776.00	3636.00	2207.00	2144.00	3731.00	3659.00
5704.00	5638.00			4194.00	4085.00
6982.00	6868.00			622.00	581.00
				3881.00	3751.00
				1240.00	1221.00
				1319.00	1284.00
				5528.00	5416.00
				7052.00	6844.00
				3881.00	3888.00
				3498.00	3426.00
				14516.00	14342.00

FEATURES

source

Location/Qualifiers
1. 188090
/Organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
/chromosome="7"

```
/map="7q31"
/clone="RP11-341C17 or IMGC:dj301"
/cell_line="Male Blood"
/clone_id="RPC-II Human Male BAC Library"
444..750
repeat_region
891..963
/xc_family="Alu"
/standard_name="SMS52952"
/ncbi_accession="U13043"
complement(2131..2409)
repeat_region
complement(8590..8831)
repeat_region
complement(10548..10834)
repeat_region
complement(11525..11795)
repeat_region
complement(18103..18403)
repeat_region
complement(27089..27369)
```

Query Match 19.4%; Score 1106; DB 9; Length 188090;

Best Local Similarity 99.9%; Pred. No. 0; Mismatches 1; Indels 0; Gaps 0;

Matches 1156; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

```
QY 1 ATGAAAGCCATGCCCCGGAAGCTGACCTGCTCTCCCACTCTCATGTGTGGGATG 60
DB 176564 ATGAAAGCCATGCCCCGGAAGCTGACCTGCTCTCCCACTCTCATGTGTGGGATG 176505
QY 61 GGCTCCCTCCCACTTGTGCTCAACCCGCGACCGCCCGCTGTCCAGAGAGAGAGAT 120
DB 176504 GGCTCCCTCCCACTTGTGCTCAACCCGCGACCGCCCGCTGTCCAGAGAGAGAT 176445
QY 121 GTCAATCTCCGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 180
DB 176444 GTCAATCTCCGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 176385
QY 181 GGAACACATTTACTTGGGGGCGCTCAATCGGATTTTAAAGCTTCCAGAGAGAGAGATC 240
DB 176384 GGAACACATTTACTTGGGGGCGCTCAATCGGATTTTAAAGCTTCCAGAGAGAGATC 176325
QY 241 TTGGTGAAGCATGAGACAGAGGCGCGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 300
DB 176324 TTGGTGAAGCATGAGACAGAGGCGCGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 176265
QY 301 GTCCAGACCTGCAATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 360
DB 176264 GTCCAGACCTGCAATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 176205
QY 361 GACTCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 420
DB 176204 GACTCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 176145
QY 421 CTGAGGCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 480
DB 176144 CTGAGGCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 176085
QY 481 TCAGGTGTCACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 540
DB 176084 TCAGGTGTCACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 176025
QY 541 GACAAAGCTGTTCAATGTCACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 600
DB 176024 GACAAAGCTGTTCAATGTCACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 175965
QY 601 AGCCGGAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 660
DB 175964 AGCCGGAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 175905
QY 661 GAGTTCTGTGGCTCGATGATTAAGTCCCTTCCAGACAGAGAGAGAGAGAGAGAGAGAGAGAG 720
DB 175904 GAGTTCTGTGGCTCGATGATTAAGTCCCTTCCAGACAGAGAGAGAGAGAGAGAGAGAGAGAG 175845
```

```
QY 721 GATATCTACTATGCTATGTTTATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 780
DB 175844 GATATCTACTATGCTATGTTTATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 175785
QY 781 CCTGAGATGAGTGTCTCCACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 840
DB 175784 CCTGAGATGAGTGTCTCCACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 175725
QY 841 GTGAGAGCTTTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 900
DB 175724 GTGAGAGCTTTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 175665
QY 901 GAGCGCAGTGGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 960
DB 175664 GAGCGCAGTGGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 175605
QY 961 GTGCTTGGCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1020
DB 175604 GTGCTTGGCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 175545
QY 1021 AAGGCGCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1080
DB 175544 AAGGCGCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 175485
QY 1081 AAGCAGATTAATGACCGGATTAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1140
DB 175484 AAGCAGATTAATGACCGGATTAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 175425
QY 1141 CTGACACTGGCGCTGAGCT 1157
DB 175424 CTGACACTGGCGCTGAGCT 175408
```

RESULT 10

AC009723 179592 bp DNA linear HTG 08-NOV-2000
LOCUS AC009723/C

DEFINITION

Home sapiens clone RP11-44F19, WORKING DRAFT SEQUENCE, 25 unordered
pieces.

ACCESSION

AC009723.4 GI:9965551

VERSION

HTG; HTGS_PHASE1; HTGS_DRAFT.

KEYWORDS

Home sapiens (human)

SOURCE

Home sapiens

ORGANISM

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

REFERENCE

1 (bases 1 to 179592)

AUTHORS

Birren,B., Linton,L., Nusbaum,C. and Lander,E.

TITLE

Home sapiens, clone RP11-44F19
Unpublished

JOURNAL

2 (bases 1 to 179592)
Birren,B., Linton,L., Nusbaum,C., Lander,E., Allen,N., Anderson,M.,
Baker,J., Baldwin,J., Barna,N., Beckert,J., Berr,M., Brown,A.,
Casper,A., Cherni,J., Collins,S., Collins,S., Collins,S., Collins,S.,
Cooke,P., DeRubeis,A., DeRubeis,A., DeRubeis,A., DeRubeis,A.,
Donelan,J., Doyle,M., Ferreira,P., Fitzhugh,N., Forrest,C.,
Funke,R., Gage,D., Galagan,J., Gargana,S., Gilbert,D., Grant,G.,
Hagos,B., Heaford,A., Horton,L., Howland,J., Jones,C., Kahn,L.,
Karatza,A., Lechoczy,J., Lieu,C., Locke,K., MacDonald,P.,
Margulis,N., McEwan,P., McGurk,A., McKernan,K., McLaughlin,J.,
Melgrim,J., Molla,M., Morris,W., Morrow,J., Mychaleckyj,J.,
Naylor,J., Nijhoff,M., O'Connor,T., O'Donnell,P., Pavlin,B.,
Peterson,K., Pollara,V., Riley,R., Roberts,D., Roy,A., Severy,P.,
Stange-Thomann,N., Stojanovic,N., Stone,C., Sudramanlan,A.,
Teasdale,S., Torkelson-Miller,T., Vassiliev,H., Vo,A., Wagner,A.,
Wheeler,J., Wu,X., Wyman,D., Ye,W.J. and Zody,M.

COMMENT

Submitted (29-AUG-1999) Whitehead Institute/MIT Center for Genome
Research, 320 Charles Street, Cambridge, MA 02141, USA
On Sep 3, 2000 this sequence version replaced gi:7770419.
All repeats were identified using RepeatMasker:
http://ftp.genome.washington.edu/RM/RepeatMasker.html
Genome Center


```

Center: Whitehead Institute/ MIT Center for Genome Research
Center code: WISR
Web site: http://www-seq.wi.mit.edu
Contact: sequence_submissions@genome.wi.mit.edu
-----
Project Information
Center project name: 11062
Center clone name: 44_F_19
-----
Summary Statistics
Sequencing vector: M13, M7815, 95% of reads
Sequencing vector: Plasmid, n/a, 5% of reads
Chemistry: Dye-Primer-amersham; 5% of reads
Chemistry: Dye-Terminator Big Dye; 95% of reads
Assembly program: Phrap; version 0.960731
Consensus quality: 159631 bases at least Q40
Consensus quality: 169781 bases at least Q30
Consensus quality: 174012 bases at least Q20
Insert size: 176000; agarose-ftp
Insert size: 177192; sum-of-contigs
Quality coverage: 4.2 in Q20 bases.
* NOTE: This is a "working draft" sequence. It currently
* consists of 25 contigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence
* as soon as it is available and the accession number will
* be preserved.
*
* 1
* 3333: contig of 3333 bp in length
*
* 3334
* 3433: gap of 100 bp
* 3434
* 4557: contig of 1124 bp in length
*
* 4558
* 4657: gap of 100 bp
* 4658
* 5689: contig of 1032 bp in length
*
* 5690
* 5790: gap of 100 bp
* 5791
* 7161: contig of 1372 bp in length
*
* 7162
* 7261: gap of 100 bp
* 7262
* 8676: contig of 1415 bp in length
*
* 8677
* 8777
* 9591: contig of 1175 bp in length
*
* 9592
* 10051: gap of 100 bp
* 10052
* 11777: contig of 1726 bp in length
*
* 11878
* 11877: gap of 100 bp
* 13012: contig of 1135 bp in length
*
* 13013
* 13112: gap of 100 bp
* 13113
* 14573: contig of 1461 bp in length
*
* 14574
* 14673: gap of 100 bp
* 14674
* 17078: contig of 2405 bp in length
*
* 17079
* 17178: gap of 100 bp
* 17179
* 18544: contig of 1366 bp in length
*
* 18545
* 18644: gap of 100 bp
* 20233: contig of 1589 bp in length
*
* 20234
* 20333: gap of 100 bp
* 20334
* 22479: contig of 214 bp in length
*
* 22480
* 22579: gap of 100 bp
* 22480
* 24247: contig of 1666 bp in length
*
* 24248
* 24347: gap of 100 bp
* 24248
* 26991: contig of 2644 bp in length
*
* 26992
* 27091: gap of 100 bp
* 27092
* 31443: contig of 4355 bp in length
*
* 31444
* 31543: gap of 100 bp
* 31544
* 37214: contig of 5675 bp in length
*
* 37215
* 37319: gap of 100 bp
* 37219
* 40338: contig of 3020 bp in length
*
* 40339
* 40438: gap of 100 bp
* 40339
* 45330: contig of 4892 bp in length
*
* 45331
* 45430: gap of 100 bp
* 45331
* 51881: contig of 7752 bp in length
*
* 51882
* 51881: gap of 100 bp
* 51883
* 61147: contig of 7865 bp in length
*
* 61148
* 61247: gap of 100 bp
* 61148
* 73146: contig of 11859 bp in length
*
* 73147
* 73247
* 82900: gap of 100 bp
* 73247
* 82900: contig of 9654 bp in length
*
* 82901
* 83000: gap of 100 bp

```

```

*      83001  139568: contig of 56566 bp in length
*      139569  139668: gap of 100 bp
*      139669  179592: contig of 39924 bp in length.
FEATURES
    source
        1..179592
            /organism="Homo sapiens"
            /mol_type="genomic DNA"
            /db_xref="taxon:9606"
            /clone="RP11-44F19"
            /clone_id="RPC1-11 Human Male BAC"
            1..3333
                /note="assembly_fragment
                clone_end:8P6
                vector_side:left"
            misc_feature
                3434..4557
                    /note="assembly_fragment"
            misc_feature
                4658..5689
                    /note="assembly_fragment"
            misc_feature
                5790..7161
                    /note="assembly_fragment"
            misc_feature
                7262..8676
                    /note="assembly_fragment"
            misc_feature
                8777..9951
                    /note="assembly_fragment"
            misc_feature
                10052..11777
                    /note="assembly_fragment"
            misc_feature
                11878..13012
                    /note="assembly_fragment"
            misc_feature
                13113..14573
                    /note="assembly_fragment"
            misc_feature
                14674..17078
                    /note="assembly_fragment"
            misc_feature
                17179..18544
                    /note="assembly_fragment"
            misc_feature
                18645..20233
                    /note="assembly_fragment"
            misc_feature
                20334..22479
                    /note="assembly_fragment"
            misc_feature
                22580..24247
                    /note="assembly_fragment"
            misc_feature
                24348..26991
                    /note="assembly_fragment"
            misc_feature
                27092..31443
                    /note="assembly_fragment"
            misc_feature
                31544..37218
                    /note="assembly_fragment"
            misc_feature
                37319..40338
                    /note="assembly_fragment"
            misc_feature
                40439..45330
                    /note="assembly_fragment"
            misc_feature
                45431..53182
                    /note="assembly_fragment"
            misc_feature
                53283..61147
                    /note="assembly_fragment"
            misc_feature
                61248..73146
                    /note="assembly_fragment"
            misc_feature
                73247..82900
                    /note="assembly_fragment"
            misc_feature
                83001..139568
                    /note="assembly_fragment"
            misc_feature
                139669..179592
                    /note="assembly_fragment
                clone_end:T7
                vector_side:right"
ORIGIN
Query Match      18.8%; Score 1068; DB 2; Length 179592;
Beet Local Similarity 99.9%; Pred. No. 0;
Matches 1188; Conservative 0; Mismatches 0; Indels 1; Gaps 1;
1 ATGAAGCCATGCCCGAAGCTGAGCTGACCTCTCTCCACCTCCATGATGAGGATG 60
97654 ATGAAGCCATGCCCGAAGCTGAGCTGACCTCTCTCCACCTCCATGATGAGGATG 97595

```

```

QY 61 GGCTCTCCATCTTGGTCAACCCGCGAGCCAGCCCGCTGTCCAGAGAGCAGGTCATT 120
Db 97594 GGCTCTCCATCTTGGTCAACCCGCGAGCCAGCCCGCTGTCCAGAGAGCAGGTCATT 97535
QY 121 GTCAATTCGAGAGAGAGCCCGCGAGGTTTCAATCAGCTGTGTGATGAGAGACA 180
Db 97534 GTCAATTCGAGAGAGAGCCCGCGAGGTTTCAATCAGCTGTGTGATGAGAGACA 97475
QY 181 GCAACATTTACTTGGGGGCGTCAATCGAATTTAACAAGCTCTCCAGCAGCTGAAGTTC 240
Db 97474 GCAACATTTACTTGGGGGCGTCAATCGAATTTAACAAGCTCTCCAGCAGCTGAAGTTC 97415
QY 241 TTGGTACGATGAGAGAGGCGCGAGACGAGAGACACCCCAAGTTTACCAACCCGCGATC 300
Db 97414 TTGGTACGATGAGAGAGGCGCGAGACGAGAGACACCCCAAGTTTACCAACCCGCGATC 97355
QY 301 GTCCAGACTGCAATGAGCCCTGACCAACCAACAAATGTCATCTCTCATTA 360
Db 97354 GTCCAGACTGCAATGAGCCCTGACCAACCAACAAATGTCATCTCTCATTA 97295
QY 361 GACTACAGAGAGAGACAGCTGATTCCTGTGGAGCCTGTGACAGGATTCGAAAGCTG 420
Db 97294 GACTACAGAGAGAGACAGCTGATTCCTGTGGAGCCTGTGACAGGATTCGAAAGCTG 97235
QY 421 CTGAGGCTGAGAGGACCTTCAAGCTGGGGAGCCTTATCATAGAAGAGCACTATCTG 480
Db 97234 CTGAGGCTGAGAGGACCTTCAAGCTGGGGAGCCTTATCATAGAAGAGCACTATCTG 97175
QY 481 TCAAGTGTCAACGAGAGCGGCTCAAGCTTGTGAAGTATGCTCTCTCAAGCAACTGAT 540
Db 97174 TCAAGTGTCAACGAGAGCGGCTCAAGCTTGTGAAGTATGCTCTCTCAAGCAACTGAT 97115
QY 541 GACAAAGCTTTCATTTGCCACGCGAGTGGAGAGCCGAGATTTTCCACCATCTCC 600
Db 97114 GACAAAGCTTTCATTTGCCACGCGAGTGGAGAGCCGAGATTTTCCACCATCTCC 97055
QY 601 AGCCGAGAACTGACCAAGAACTCTGAGCGGATGAGCAATGTCGCTTCACTTCATGAT 660
Db 97054 AGCCGAGAACTGACCAAGAACTCTGAGCGGATGAGCAATGTCGCTTCACTTCATGAT 96995
QY 661 GAGTTCGTCCTGATGATTAATATCCCTTC - GAGACCTTCAACCATCTCCATCTT 719
Db 96994 GAGTTCGTCCTGATGATTAATATCCCTTCGAGACCTTCAACCATCTCCATCTT 96935
QY 720 TGATATCTACTATGATGATTTTGAAGAGTGCACCTTTGCTTATCTTTGAACCTTCA 779
Db 96934 TGATATCTACTATGATGATTTTGAAGAGTGCACCTTTGCTTATCTTTGAACCTTCA 96875
QY 780 ACTTGAGATGATGATTCACCAAGGCTTCAACCAAGAGAGAGTGTATATCTCAAGCT 839
Db 96874 ACTTGAGATGATGATTCACCAAGGCTTCAACCAAGAGAGAGTGTATATCTCAAGCT 96815
QY 840 CGTGAAGTTTGAAGAGAGACAGAGCTTCACTCTATGATGAGTGCACCTTTGCTG 899
Db 96814 CGTGAAGTTTGAAGAGAGACAGAGCTTCACTCTATGATGAGTGCACCTTTGCTG 96755
QY 900 TGAAGGCAAGTGGGATGAGTACCGCTGTGTCAGAGCTGCTTCTGTCCAAAGGGGGGC 959
Db 96754 TGAAGGCAAGTGGGATGAGTACCGCTGTGTCAGAGCTGCTTCTGTCCAAAGGGGGGC 96695
QY 960 CGTGCTTGGCAGAGCCCTTGAAGTCACTCCAGATGATGATGATGATGATGATGATGAT 1019
Db 96694 CGTGCTTGGCAGAGCCCTTGAAGTCACTCCAGATGATGATGATGATGATGATGATGAT 96635
QY 1020 CAAAGGCGAGAGCGGAAATGAATCCCTGATGATGATGATGATGATGATGATGATGAT 1079
Db 96634 CAAAGGCGAGAGCGGAAATGAATCCCTGATGATGATGATGATGATGATGATGATGAT 96575
QY 1080 GAAAGATTAATAGCGGATTAAGAGAGCGGCTGAGTCTTGTTAACGGGGCGAGGGCAC 1139
Db 96574 GAAAGATTAATAGCGGATTAAGAGAGCGGCTGAGTCTTGTTAACGGGGCGAGGGCAC 96515

```

```

QY 1140 GCTGACCTGCGCTGCTCAAGTGAAGAGACATCCCTGACAGAGTSCG 1188
Db 96514 GCTGACCTGCGCTGCTCAAGTGAAGAGACATCCCTGACAGAGTSCG 96466

RESULT 11
HSM805617
LOCUS
DEFINITION
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.
1 (bases, 1 to 3556)
Ostenwaeider, B., Obermaier, B., Newes, H. W., Gaassenhuber, J. and
Wiemann, S.
Direct Submission
Submitted (09-JUL-2002) 1, D-85764 Neuherberg, GERMANY
Clone from S. Wiemann, Molecular Genome Analysis, German Cancer
Research Center (DKFZ); Email: s.wiemann@dkfz-heidelberg.de;
sequenced by MedGenomix (Martinried/Germany) within the CDNA
sequencing consortium of the German Genome Project. This clone
(DKFZp56600546) is available at the RZPD in Berlin. Please contact
the RZPD: Ressourcenzentrum, Heubnerweg 6, 14059
Berlin-Charlottenburg, GERMANY; Email: clone@rzpd.de Further
information about the clone and the sequencing project is available
at http://mips.gsf.de/proj/cdna/.
Location/Qualifiers
1..3556
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="RZPD:DKFZp56600546"
/db_xref="taxon:9606"
/clone="DKFZp56600546"
/tissue_type="kidney"
/clone_lib="566 (synonym: hfk2). Vector pAMP1; host
X1-2blue; sites NotI + SalI"
/dev_stage="fetal"
1..3556
/gene="DKFZp56600546"
<1..1031
/gene="DKFZp56600546"
/feature="strong similarity to mouse plexin 2"
/codon_start=3
/product="hypothetical protein"
/protein_id="CAD39161.1"
/db_xref="GI:21740305"
/db_xref="SPTREMBL:Q8ND00"
/translation="LEWROGSGAGATLLODEDITIKIENDMKRLNTLAHYVDPGSVVA
LVSKQVAVNAVNNSVTRTSASKYEMNIRRTGSPDSRSTPMITPLTESGVMMWL
VKNHEHGOQKEDRSKSNVSIYLRLATIGTLOKFDVDDIFETIFSAHSGALPLA
IKYNEHFLDEQADKRGHDPVRRHWNKNCPLRPFWMDLNQPOVDFHKSITDLC
LSVAQTFMDSCTSEHRLGKDSPSNKLIVKDVPSYKNWERYSDIGKMPALSDOD
KNAYLAEOSRHNMNEFNMTSALSEIFSVYKXSEIILGLDHDQCGKQKLAYKLEQV
ITIMSLDS"
3532
/gene="DKFZp56600546"

polyA_site

ORIGIN
Query Match 15.4%; Score 878; DB 9; Length 3556;
Best Local Similarity 99.7%; Pred. No. 0;
Matches 1028; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 4655 ATCTGAGTGGGAGACAGAAATGTTGGGCAAGATGATCTTGAAGATGAAGACATCACA 4714
Db 1 ATCTGAGTGGGAGACAGAAATGTTGGGCAAGATGATCTTGAAGATGAAGACATCACA 60
QY 4715 CCAGATGAGATGATTTGAAGAGCACTGAGCAGCTGAGCAGCTGAGCAGCTGAGCAGCTG 4774
Db 61 CCAGATGAGATGATTTGAAGAGCACTGAGCAGCTGAGCAGCTGAGCAGCTGAGCAGCTG 120

```

4775 GTTCCGTTGGGATAGTGTGCAAGGATGACAGGCTATTAACGAGTGAACAATCTCA 4834
Db 121 GTTCCGTTGGGATAGTGTGCAAGGATGACAGGCTATTAACGAGTGAACAATCTCA 180
Qy 4835 CCGTCTCCAGAGCTCCAGCAAGTAAATATGAAAAATGATCCGGTACACGGGACGCCCG 4894
Db 181 CCGTCTCCAGAGCTCCAGCAAGTAAATATGAAAAATGATCCGGTACACGGGACGCCCG 240
Qy 4895 AACAGCTCCGCTCAGAGACACCTATGATCACTCTCTGACCTGGAGAGTGAAGTCAATG 4954
Db 241 AACAGCTCCGCTCAGAGACACCTATGATCACTCTCTGACCTGGAGAGTGAAGTCAATG 300
Qy 4955 GGCACCTAGTGAAGAACCAAGAGCAGGAGACAGAAAGAGAGGAGGAGCCGGGAGAGCA 5014
Db 301 GGCACCTAGTGAAGAACCAAGAGCAGGAGACAGAAAGAGAGGAGGAGCCGGGAGAGCA 360
Qy 5015 TGGTGTCTGAATCTTACCTGACCCGACTCTCTGGGCACTAAGGGGACACTGCAAGATTG 5074
Db 361 TGGTGTCTGAATCTTACCTGACCCGACTCTCTGGGCACTAAGGGGACACTGCAAGATTG 420
Qy 5075 TGGATGACCTCTTGAAGACCATCTTTCAGACAAGGACACGTTGGCTCTGCCCTCCCTG 5134
Db 421 TGGATGACCTCTTGAAGACCATCTTTCAGACAAGGACACGTTGGCTCTGCCCTCCCTG 480
Qy 5135 CCATCAAGTACATGTTTGAATCTTCTGGATGAGCAGGCTGATTAACATGGCATTCATGAC 5194
Db 481 CCATCAAGTACATGTTTGAATCTTCTGGATGAGCAGGCTGATTAACATGGCATTCATGAC 540
Qy 5195 CGCAGCTCCGCAATACCTGGAAGAGCAATGCTGCCCCCTGAGAGTTTGGGTCAACATGA 5254
Db 541 CGCAGCTCCGCAATACCTGGAAGAGCAATGCTGCCCCCTGAGAGTTTGGGTCAACATGA 600
Qy 5255 TCGAAGACCCGCAATGTTGTGTTTGAATCCATTAAGAACAGATACAGACGCTGCTCT 5314
Db 601 TCGAAGACCCGCAATGTTGTGTTTGAATCCATTAAGAACAGATACAGACGCTGCTCT 660
Qy 5315 CTGATGAGGCTCAGACCTTCATGAGACTCTTCTCCAGCTCAGACACCGGCTGGGCAAG 5374
Db 661 CTGATGAGGCTCAGACCTTCATGAGACTCTTCTCCAGCTCAGACACCGGCTGGGCAAG 720
Qy 5375 ACTGCGCTCCAAACAAGCTGTGTATGCAAGACATCCCAAGCTTCAAGAAATTGGGTG 5434
Db 721 ACTGCGCTCCAAACAAGCTGTGTATGCAAGACATCCCAAGCTTCAAGAAATTGGGTG 780
Qy 5435 AGAGATTATCTCAGACATAGGAGAAATGCGAGCATACAGGACCAAGACATGAAAGCAT 5494
Db 781 AGAGATTATCTCAGACATAGGAGAAATGCGAGCATACAGGACCAAGACATGAAAGCAT 840
Qy 5495 ACCTGCTGAGACAGTCCCGAGTGCATGAAATGAGTTCAACACATAGTGCATCTCAG 5554
Db 841 ACCTGCTGAGACAGTCCCGAGTGCATGAAATGAGTTCAACACATAGTGCATCTCAG 900
Qy 5555 AGATTTCTCTCTAGTGGGCAAAATACAGGAGAGAGATCTTGGACCTCTGACCAAGATG 5614
Db 901 AGATTTCTCTCTAGTGGGCAAAATACAGGAGAGAGATCTTGGACCTCTGACCAAGATG 960
Qy 5615 ACCAGTGTGGAGAGCAAACTGGCTTCAAACTAGAAACAAGTCAATACCTCATAGACT 5674
Db 961 ACCAGTGTGGAGAGCAAACTGGCTTCAAACTAGAAACAAGTCAATACCTCATAGACT 1020
Qy 5675 TAGACAGCTGA 5685
Db 1021 TAGACAGCTGA 1031

RESULT 12
AX574564
LOCUS AX574564 2597 bp DNA 1linear PAT 07-JAN-2003
DEFINITION Sequence 91 from Patent WO2224888.
ACCESSION AX574564
VERSION AX574564.1 GI:27551835
KEYWORDS

SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE 1
AUTHORS Baker, K.P., Eaton, D.L., Filvaroff, E., Goddard, A., Grimaldi, J.C.,
Gurney, A.L., Smith, V., Stephan, D.P., Matanabe, C.K., Wood, W.I.,
Zhang, Z. and Fong, S.
TITLE Secreted and transmembrane polypeptides and nucleic acids encoding
the same
JOURNAL Patent: WO 0224888-A 91 28-MAR-2002;
GENENTECH, INC. (US)
FEATURES
source
1. 2597
Location/Qualifiers
/organism="Homo sapiens"
/mol_type="unassigned DNA"
/db_xref="taxon:9606"
ORIGIN
Query Match 13.9%; Score 793; DB 6; Length 2597;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 793; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 579 CGAGTATTTTCCCAACATCTCCAGCCGGAACCTGACCAAGAACTCTGAGCGGATGAGCAT 638
Db 1 CGAGTATTTTCCCAACATCTCCAGCCGGAACCTGACCAAGAACTCTGAGCGGATGAGCAT 60
Qy 639 GTTGGCGTACGCTCTTCCATGATGATGATGATGATGATGATGATGATGATGATGATG 698
Db 61 GTTGGCGTACGCTCTTCCATGATGATGATGATGATGATGATGATGATGATGATGATG 120
Qy 699 CTTACACATCAATCCCTGACCTTGTATATCTAATGATGATGATGATGATGATGATGAT 758
Db 121 CTTACACATCAATCCCTGACCTTGTATATCTAATGATGATGATGATGATGATGATGAT 180
Qy 759 TGTCTACTTTTGTGACCTCTCCAACTGAGATGATGATGATGATGATGATGATGATGAT 818
Db 181 TGTCTACTTTTGTGACCTCTCCAACTGAGATGATGATGATGATGATGATGATGATGAT 240
Qy 819 GGAAGGTATATACATCAAGCTCGTGAAGCTTTGCAAGAGGACACAGGCTTCACTCTCT 878
Db 241 GGAAGGTATATACATCAAGCTCGTGAAGCTTTGCAAGAGGACACAGGCTTCACTCTCT 300
Qy 879 TGTAAAGTGTCCATGCTGTGAGCGAGTGGGATGAGTACCGCTCTGTGAGGCTGC 938
Db 301 TGTAAAGTGTCCATGCTGTGAGCGAGTGGGATGAGTACCGCTCTGTGAGGCTGC 360
Qy 939 CTACCTGTCCAAAGGGGGGGGGCGGTCTTGGAGGACCTTGGAGTGCATCCAGATGATGA 998
Db 361 CTACCTGTCCAAAGGGGGGGGGCGGTCTTGGAGGACCTTGGAGTGCATCCAGATGATGA 420
Qy 999 CCTGTCTTCAACCGCTCTTCTCCAAAGGGGCAAGAGGCAAAATGAATCTCTGATGATC 1058
Db 421 CCTGTCTTCAACCGCTCTTCTCCAAAGGGGCAAGAGGCAAAATGAATCTCTGATGATC 480
Qy 1059 GGCCTGTGCAATCTTCAATCTTGAAGCAAGATTAAGACCGCATTAAGAGCGGGCTGATC 1118
Db 481 GGCCTGTGCAATCTTCAATCTTGAAGCAAGATTAAGACCGCATTAAGAGCGGGCTGATC 540
Qy 1119 TTGTTACCGGGGCGAGGCGACGCTGACCTGAGCTGCTCAAGTGAAGGACATCCCTG 1178
Db 541 TTGTTACCGGGGCGAGGCGACGCTGACCTGAGCTGCTCAAGTGAAGGACATCCCTG 600
Qy 1179 CAGCAGTGGCTCTTAACTTGAAGCAATCTTCTGAGGCTGAGCAATGATGCTCCCT 1238
Db 601 CAGCAGTGGCTCTTAACTTGAAGCAATCTTCTGAGGCTGAGCAATGATGCTCCCT 660
Qy 1239 GGAAGTGTCCGACATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1298
Db 661 GGAAGTGTCCGACATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 720
Qy 1299 GTCTGTATCGATATGCTTACAAAGACATCTCTTGGCTTTTGGGGCACCAAAAGTGG 1358

Db 721 GTCTGTCATGCATATGCTACAGAACCACTCTCTGCGCTTTGTGGGACCAAAAGTGG 780

Qy 1359 CAAGCTGAAGAG 1371

Db 781 CAAGCTGAAGAG 793

RESULT 13

LOCUS AY358850 2597 bp mRNA linear PRI 03-OCT-2003

DEFINITION Homo sapiens clone DNA190803 FAYV2820 (UNQ2820) mRNA, complete cds.

ACCESSION AY358850

VERSION AY358850.1 GI:37182817

KEYWORDS FLI CDNA.

SOURCE Homo sapiens (human)

ORGANISM Homo sapiens

REFERENCE 1 (bases 1 to 2597)

AUTHORS Clark,H.F., Gurney,A.L., Abaya,E., Baker,K., Baldwin,D., Bush,J., Chen,J., Chow,B., Chui,C., Crowley,C., Currell,B., Deuel,B., Dowd,P., Eaton,D., Foster,J., Grimaldi,C., Gu,Q., Hass,P.E., Heidens,S., Huang,A., Kim,H.S., Klimowski,L., Jin,Y., Johnson,S., Lee,J., Lewis,L., Liao,D., Mark,M., Robbie,E., Sanchez,C., Schoenfeld,J., Sehnagiri,S., Simmons,L., Singh,J., Smith,V., Stinson,J., Vagter,A., Vandlen,R., Watanabe,C., Wieland,D., Woods,K., Xie,M.H., Yansura,D., Yi,S., Yu,G., Yuan,J., Zhang,M., Zhang,Z., Goddard,A., Wood,W.I. and Godowski,P.

THE SECRETED PROTEIN DISCOVERY INITIATIVE (SPDI), A LARGE-SCALE EFFORT TO IDENTIFY NOVEL HUMAN SECRETED AND TRANSMEMBRANE PROTEINS: A BIOINFORMATIC ASSESSMENT

JOURNAL Genome Res. 13 (10), 2265-2270 (2003)

PubMed 12975309

REFERENCE 2 (bases 1 to 2597)

AUTHORS Clark,H.F.

TITLE Direct Submission

JOURNAL Submitted (01-AUG-2003) Department of Bioinformatics, Genentech, Inc., 1 DNA Way, South San Francisco, CA 94080, USA

FEATURES

source

1..2597

/organism="Homo sapiens"

/mol_type="mRNA"

/db_xref="taxon:9606"

/clone="DNA190803"

1..2597

/locus_tag="UNQ2820"

59..901

/locus_tag="UNQ2820"

/note="PRO34003"

/codon_start=1

/product="FAYV2820"

/protein_id="AA089209.1"

/db_xref="GI:37182818"

/translation="MFAYVHDFVAMNIRKIPSTFTTTPIDYIYVGGSGFYF LTLQEMWSPSTSTKEQVYTKVLCKEPTANSVEVPIGEGSGVERLLQAY LSKAGAVLGRITGVHDDLLFTFVSKGQKMKSLDESALCIPIILQINDR KERQ SCYREGTLDLMLKVDI PCSALLITIDNFCGLDMNAPAGVSDMVRGIPVTEEDR RMTSIVAVVYKNSLAFVGTSGKLVKVPVSLCPTLELQTPRSHRATVTEILFSS CSSN"

ORIGIN

Query Match 13.9%; Score 793; DB 9; Length 2597;

Best Local Similarity 100.0%; Pred. No. 0;

Matches 793; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 579 CGAGATTTTCCACCATCTCCAGCGGAAATGACCAAGACTCTGAGCGAGTGCAT 638

Db 1 CGAGATTTTCCACCATCTCCAGCGGAAATGACCAAGACTCTGAGCGAGTGCAT 60

Qy 639 GTTCGCGTACGCTTTCATGATGAGTGTGCGCTCGATGATTAAGATCCCTTGACAC 698

Db 61 GTTCGCGTACGCTTTCATGATGAGTGTGCGCTCGATGATTAAGATCCCTTGACAC 120

Qy 699 CTTCACATCATCCCTGACTTGATATCTATGTTTGTAGCAGTGGCACTT 758

Db 121 CTTCACATCATCCCTGACTTGATATCTATGTTTGTAGCAGTGGCACTT 180

Qy 759 TGTCTACTTTTGAACCTTCACACCTGAGATGTGTCTCACAGGCTCCACCAAGGA 818

Db 181 TGTCTACTTTTGAACCTTCACACCTGAGATGTGTCTCACAGGCTCCACCAAGGA 240

Qy 819 GCAGGTATATACATCAAGCTGATGAGCTTTGCAAGAGACACAGCTTCACTCTTA 878

Db 241 GCAGGTATATATACATCAAGCTGATGAGCTTTGCAAGAGACACAGCTTCACTCTTA 300

Qy 879 TGTAGAGTGGCCCATTTGGCTGTGACCGCAGATGGGGGTGAGTACCCCTGCTGACAGCTGC 938

Db 301 TGTAGAGTGGCCCATTTGGCTGTGACCGCAGATGGGGGTGAGTACCCCTGCTGACAGCTGC 360

Qy 939 CTACCTGTCCAAAGCGGGGCGGTGCTTGACAGACCTTGAGATCCATCCAGATGATGA 998

Db 361 CTACCTGTCCAAAGCGGGGCGGTGCTTGACAGACCTTGAGATCCATCCAGATGATGA 420

Qy 999 CCTGCTTTACCGTCTTCTCCAGAGGCGCAGAGCGGAAATGAAATCCCTGATGAGTGC 1058

Db 421 CCTGCTTTACCGTCTTCTCCAGAGGCGCAGAGCGGAAATGAAATCCCTGATGAGTGC 480

Qy 1059 GGCCCTGTGACATCTTCACTTGAAGCAGTAAATGACCGCATTAAGAGCGGCTGACATC 1118

Db 481 GGCCCTGTGACATCTTCACTTGAAGCAGTAAATGACCGCATTAAGAGCGGCTGACATC 540

Qy 1119 TTGTTACCGGGGCGAGGGGACGCTGACCTGGCTGAGTCAAGTGAAGACATCCCTG 1178

Db 541 TTGTTACCGGGGCGAGGGGACGCTGACCTGGCTGAGTGAAGACATCCCTG 600

Qy 1179 CAGCAGTGGCTCTTACCATTTGACATTAATCTTGTGGCTGGCATGAATGCTCCCT 1238

Db 601 CAGCAGTGGCTCTTACCATTTGACATTAATCTTGTGGCTGGCATGAATGCTCCCT 660

Qy 1239 GGGAGTGTCCGACATGATGCGTGAATTCCTCTTTCACGAGACAGGAGCCGATGAC 1298

Db 661 GGGAGTGTCCGACATGATGCGTGAATTCCTCTTTCACGAGACAGGAGCCGATGAC 720

Qy 1299 GTCTGTCATGCATATGCTTACAGAACCACTCTTGCGCTTTGGGACCAAAAGTGG 1358

Db 721 GTCTGTCATGCATATGCTTACAGAACCACTCTTGCGCTTTGGGACCAAAAGTGG 780

Qy 1359 CAAGCTGAAGAG 1371

Db 781 CAAGCTGAAGAG 793

RESULT 14

LOCUS HSM802037 1029 bp mRNA linear PRI 18-FEB-2000

DEFINITION Homo sapiens mRNA, cDNA DKFZp434G0625 (from clone DKFZp434G0625); partial cds.

ACCESSION AL137352

VERSION AL137352.1 GI:6807867

KEYWORDS

SOURCE Homo sapiens (human)

ORGANISM Homo sapiens

REFERENCE 1 (bases 1 to 1029)

AUTHORS Amstutz,W., Winkler,U., Mewes,H.W., Gassenhuber,J. and Wiemann,S.

TITLE Direct Submission

JOURNAL Submitted (15-JAN-2000) MIPS, Am Klopferstr. 18a, D-82152 Martinsried, GERMANY

COMMENT Clone from S. Wiemann, Molecular Genome Analysis, German Cancer Research Center (DKFZ), Email s.wiemann@kfz-heidelberg.de; sequenced by EMBL (European Molecular Biology Laboratories, Heidelberg/Germany) within the cDNA sequencing consortium of the German Genome Project.

This clone (DKFZp434G0625) is available at the RZPD in Berlin.

Please contact the RZPD: Ressourcenzentrum, Heubnerweg 6, 14059

Berlin-Charlottenburg, GERMANY; Email: clone@zpd.de Further information about the clone and the sequencing project is available at <http://www.mips.biochem.mpg.de/proj/cdn/>.

FEATURES

SOURCE

1.1029
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /clone="DKFZp434G0625"
 /tissue_type="testis"
 /clone_id="434 (synonym: htes3). Vector pSPORT1; host DH10B; sites NotI + SalI"
 /dev_stage="adult"
 1.1029
 /gene="DKFZp434G0625"
 /gene="DKFZp434G0625"
 /note="strong similarity to plexins"
 /codon_start=2
 /product="hypothetical protein"
 /protein_id="CAI70707.1"
 /db_xref="GI:6807868"
 /db_xref="SPRIMBL:Q9NTD4"
 /translation="KNSBADGMFAVYHDEFVASMIRIADFTLLIPDPIYYVGFSGNFVPLTQPEVSPSPGTTKEQVTSKLVKCKEDTAFNSVVEVPIGSGVEYRLDPAVTSKAGNTLGRITGLVHPDDLLFTVSGQRKMSLDSEALCTFIKKINDRIKERLOSCTRGESTLDLAWKVKDIPSSALTLTDNFGGLDMNAPLGVSDNVRGIPVFTEDRDMTSVIAVYVKNHSLAVFGTSGLKFSFGTGGGIGTDEWIVESDPPGA NIASQBMFLCVTLQCSHKAISDORVPLLCFFLVNFGNS"
 991.396
 /gene="DKFZp434G0625"
 1006
 /gene="DKFZp434G0625"

ORIGIN

Query Match 12.4%; Score 706; DB 9; Length 1029;
 Best Local Similarity 99.9%; Pred. No. 0;
 Matches 756; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

615 CAAGAAGCTGAGGGGATGAGTGGCGTACGTCCTCATAGATGATTCGAGGCTC 674
 1 CAAGAAGCTGAGGGGATGAGTGGCGTACGTCCTCATAGATGATTCGAGGCTC 60
 675 GATGATTAAGATCCCTTCGACACCTTCACCATCATCCCTGATTTGATATCTAATGT 734
 61 GATGATTAAGATCCCTTCGACACCTTCACCATCATCCCTGATTTGATATCTAATGT 120
 735 CATAGCTTTAGCAATGCACTTTGATCTTTTGAACCTTCAACCTGAGATGAGTGC 794
 121 CATAGCTTTAGCAATGCACTTTGATCTTTTGAACCTTCAACCTGAGATGAGTGC 180
 735 TCACCAAGGCTCCACCAAGGAGAGAGATGATATCATCAAGCTGAGAGCTTTTCAA 854
 181 TCACCAAGGCTCCACCAAGGAGAGAGATGATATCATCAAGCTGAGAGCTTTTCAA 240
 855 GAGAGACACAGGCTTCAACTCTTATGTAAGAGTGGCCATTTGGCTGTAGCGCAATGGGCT 914
 241 GAGAGACACAGGCTTCAACTCTTATGTAAGAGTGGCCATTTGGCTGTAGCGCAATGGGCT 300
 915 GAGATACGCGCTGCGAGAGGCTGCTTACTGTCAAACGGGGGCGGTGCTGCGAGAGAC 974
 301 GAGATACGCGCTGCGAGAGGCTGCTTACTGTCAAACGGGGGCGGTGCTGCGAGAGAC 360
 975 CTTGAGAGTCATCCAGATGATGACCTGCTTCAACCGTCTTTCTCAAGGCGCAAGACG 1034
 361 CTTGAGAGTCATCCAGATGATGACCTGCTTCAACCGTCTTTCTCAAGGCGCAAGACG 420
 1035 GAAATGAATCCCTGATGATGAGCGCCCTGTCATCTTCAATCTTGAAGCAGATTAATGA 1094
 421 GAAATGAATCCCTGATGATGAGCGCCCTGTCATCTTCAATCTTGAAGCAGATTAATGA 480
 1095 CCGCATTAAGAGAGCGGCTGAGCTTTGTTACCGGGGCGAGGACGCTGACCTGGGCTG 1154

Db 481 CCGCATTAAGAGAGCGGCTGAGCTTTGTTACCGGGGCGAGGACGCTGAGCTGGGCTG 540
 QY 1155 GCTCAAGGTGAAGACATCCCTTCGAGAGAGTGGCTTTAAACATTTGACATTAATCTTG 1214
 Db 541 GCTTAAGGTGAAGACATCCCTTCGAGAGAGTGGCTTTAAACATTTGACATTAATCTTG 600
 QY 1215 TGGCTTGAAGTGAATGCTTCCCTTGGAGTGTGTCGACATTTGGTGGTGAATCCCGTCT 1274
 Db 601 TGGCTTGAAGTGAATGCTTCCCTTGGAGTGTGTCGACATTTGGTGGTGAATTTCCGCTT 660
 QY 1275 CACGAGAGACAGGAGCCCATGACCTCTGTCATGCGATATGCTTACAGAACCATCTCT 1334
 Db 661 CACGAGAGACAGGAGCCCATGACCTCTGTCATGCGATATGCTTACAGAACCATCTCT 720
 QY 1335 GGCCTTTTGGGACACAAAAGTGGCAAGCTGAAG 1371
 Db 721 GGCCTTTTGGGACACAAAAGTGGCAAGCTGAAG 757

RESULT 15

AC025595/c 74655 bp DNA linear HTG 26-JUL-2000
 LOCUS AC025595
 DEFINITION Homo sapiens clone RP11-24E10, LOW-PASS SEQUENCE SAMPLING.
 ACCESSION AC025595
 VERSION AC025595.2 GI:7341984
 KEYWORDS HTG; HTGS PHASE0.
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

REFERENCE 1 (bases 1 to 74655)
 Birren, B., Linton, L., Nussbaum, C. and Lander, E.
 Homo sapiens, clone RP11-24E10
 Unpublished
 2 (bases 1 to 74655)
 Birren, B., Linton, L., Nussbaum, C., Lander, E., Abraham, H., Allen, N., Anderson, S., Baldwin, J., Barina, N., Bastien, V., Bedalov, F., Boguslavsky, I., Boukhalter, B., Brown, A., Burkett, G., Campodolito, A., Castle, A., Choepel, Y., Colangelo, M., Collins, S., Collymore, A., Cooke, P., Dearrell, K., Dewar, K., Diaz, J. S., Dodge, S., Domingo, M., Doyle, M., Ferreira, P., Fitzhugh, W., Gage, D., Galsagan, J., Gardyna, S., Ginde, S., Goyle, M., Graham, L., Grand-Pierre, N., Grant, G., Hagos, B., Heaford, A., Horton, L., Howland, J. C., Iliev, I., Johnson, R., Jones, C., Kam, L., Karakas, A., Klein, J., Laroque, K., Lamazares, R., Lander, E., Lehoczy, J., Levine, R., Lieu, C., Liu, G., Locke, K., MacDonald, P., Margulis, N., McCarthy, M., McEwan, P., McGuirk, A., McKernan, K., McPheters, R., Meldrum, J., Meneses, L., Mihova, T., Miranda, C., Mlenka, P., Morrow, J., Murphy, T., Naylor, J., Norman, C. H., O'Connor, T., O'Donnell, P., O'Neill, D., Oliver, T. M., Oliver, J., Peterson, K., Pierre, N., Pisan, C., Pollara, V., Raymond, C., Riley, R., Rogov, P., Rothman, D., Roy, A., Santos, R., Schauer, S., Severy, P., Spencer, B., Stange, Thomann, N., Stojanovic, N., Subramanian, A., Talamas, J., Teste, S., Theodore, J., Tittel, A., Travers, M., Triggillo, J., Vassiliev, H., Viel, R., Vo, A., Wilson, B., Wu, X., Wyman, D., Ye, W. J., Young, G., Zainoun, J., Zimmer, A. and Zody, M.

TITLE
 JOURNAL

COMMENT

Submitted (12-MAR-2000) Whitehead Institute/MIT Center for Genome Research, 320 Charles Street, Cambridge, MA 02141, USA
 On Mar 30, 2000 this sequence version replaced gi:7229794.
 All repeats were identified using RepeatMasker:
 Smith, A.F.A. & Green, P. (1996-1997)
<http://ftp.genome.washington.edu/RM/RepeatMasker.html>

----- Genome Center

Center: Whitehead Institute/ MIT Center for Genome Research

Center code: WIBR

Web site: <http://www-seq.wi.mit.edu>

Contact: sequence_submissions@genome.wi.mit.edu

----- Project Information

Center Project name: L4440

Center clone name: 24_E_10

* NOTE: This record contains 90 individual

* sequencing reads that have not been assembled into
* contigs. Runs of N are used to separate the reads
* and the order in which they appear is completely
* arbitrary. Low-pass sequence sampling is useful for
* identifying clones that may be gene-rich and allows
* overlap relationships among clones to be deduced.
* However, it should not be assumed that this clone
* will be sequenced to completion. In the event that
* the record is updated, the accession number will
* be preserved.

1 727: contig of 727 bp in length
728 827: gap of 100 bp
828 1590: contig of 763 bp in length
1591 1690: gap of 100 bp
1691 2443: contig of 753 bp in length
2444 2543: gap of 100 bp
2544 3286: contig of 743 bp in length
3287 3386: gap of 100 bp
3387 4099: contig of 713 bp in length
4100 4199: gap of 100 bp
4200 4924: contig of 725 bp in length
4925 5024: gap of 100 bp
5025 5745: contig of 721 bp in length
5746 5846: gap of 100 bp
5846 6567: contig of 722 bp in length
6567 6667: gap of 100 bp
6668 7409: contig of 742 bp in length
7410 7509: gap of 100 bp
7510 8339: contig of 730 bp in length
8340 9048: gap of 100 bp
9049 9149: gap of 100 bp
9149 9902: contig of 754 bp in length
9903 10002: gap of 100 bp
10003 10745: contig of 743 bp in length
10746 10845: gap of 100 bp
10846 11593: contig of 748 bp in length
11594 11693: gap of 100 bp
11694 12426: contig of 733 bp in length
12427 12526: gap of 100 bp
12527 13240: contig of 714 bp in length
13241 13341: gap of 100 bp
13341 14074: contig of 734 bp in length
14075 14174: gap of 100 bp
14175 14806: contig of 732 bp in length
14807 15006: gap of 100 bp
15007 15128: contig of 722 bp in length
15129 15828: gap of 100 bp
15829 16542: contig of 714 bp in length
16543 16642: gap of 100 bp
16643 17379: contig of 737 bp in length
17380 17479: gap of 100 bp
17480 18325: contig of 746 bp in length
18326 18325: gap of 100 bp
18326 19056: contig of 731 bp in length
19057 19156: gap of 100 bp
19157 19894: contig of 738 bp in length
19895 19994: gap of 100 bp
19995 20729: contig of 735 bp in length
20730 20829: gap of 100 bp
20830 21569: contig of 740 bp in length
21570 21669: gap of 100 bp
21670 22401: contig of 732 bp in length
22402 22501: gap of 100 bp
22502 23325: contig of 724 bp in length
23326 23325: gap of 100 bp
23326 24053: contig of 728 bp in length
24054 24153: gap of 100 bp
24154 24892: contig of 739 bp in length
24893 24992: gap of 100 bp
24993 25727: contig of 735 bp in length
25728 25827: gap of 100 bp
25829 26566: contig of 739 bp in length

26567 26666: gap of 100 bp
26667 27110: contig of 744 bp in length
27111 27510: gap of 100 bp
27511 28341: contig of 731 bp in length
28342 28341: gap of 100 bp
28342 29090: contig of 749 bp in length
29091 29190: gap of 100 bp
29191 29923: contig of 733 bp in length
29924 30023: gap of 100 bp
30024 30742: contig of 719 bp in length
30743 30842: gap of 100 bp
30843 31558: contig of 716 bp in length
31559 31558: gap of 100 bp
31560 32388: contig of 730 bp in length
32389 32488: gap of 100 bp
32489 33204: contig of 716 bp in length
33205 33304: gap of 100 bp
33305 34018: contig of 714 bp in length
34019 34118: gap of 100 bp
34119 34850: contig of 733 bp in length
34851 34950: gap of 100 bp
34951 35689: contig of 733 bp in length
35690 35789: gap of 100 bp
35790 36336: contig of 747 bp in length
36337 36636: gap of 100 bp
36637 37354: contig of 718 bp in length
37355 37454: gap of 100 bp
37455 38198: contig of 744 bp in length
38199 38298: gap of 100 bp
38299 39045: contig of 747 bp in length
39046 39145: gap of 100 bp
39146 39872: contig of 727 bp in length
39873 39972: gap of 100 bp
39973 40679: contig of 707 bp in length
40680 40779: gap of 100 bp
40780 41511: contig of 732 bp in length
41512 41611: gap of 100 bp
41612 42342: contig of 731 bp in length
42343 42442: gap of 100 bp
42443 43178: contig of 736 bp in length
43179 43278: gap of 100 bp
43279 44022: contig of 744 bp in length
44023 44122: gap of 100 bp
44123 44860: contig of 738 bp in length
44861 44960: gap of 100 bp
44961 45676: contig of 716 bp in length
45677 45776: gap of 100 bp
45777 46515: contig of 733 bp in length
46516 46615: gap of 100 bp
46616 47350: contig of 735 bp in length
47351 47450: gap of 100 bp
47451 48148: contig of 698 bp in length
48149 48248: gap of 100 bp
48249 48949: contig of 701 bp in length
48950 49049: gap of 100 bp
49050 49775: contig of 726 bp in length
49776 49875: gap of 100 bp
49876 50592: contig of 717 bp in length
50593 51419: contig of 727 bp in length
51420 51519: gap of 100 bp
51520 52272: contig of 753 bp in length
52273 52372: gap of 100 bp
52373 53123: contig of 751 bp in length
53124 53223: gap of 100 bp
53224 53965: contig of 742 bp in length
53966 54065: gap of 100 bp
54066 54780: contig of 715 bp in length
54781 54880: gap of 100 bp
54881 55621: contig of 741 bp in length
55622 55721: gap of 100 bp
55722 56458: contig of 737 bp in length
56459 56558: gap of 100 bp

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: February 19, 2004, 23:14:05 ; Search time 1340 seconds
(without alignments)
18042.161 Million cell updates/sec

Title: US-09-964-956-12
Perfect score: 5691
Sequence: 1 atgaagccatgcctcgaa.....gcttagacagctgaataaa 5691

Scoring table: OLIGO_NTIC
Gapop 60.0 , Gapext 60.0

Searched: 3373863 seqs, 212409041 residues

Word size : 0
Total number of hits satisfying chosen parameters: 6747726

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Listing first 1000 summaries

Database : N_Geneseq_293an04:*
1: Geneseq19808:*
2: Geneseq19908:*
3: Geneseq20008:*
4: Geneseq20018:*
5: Geneseq20028:*
6: Geneseq20038:*
7: Geneseq20048:*
8: Geneseq20058:*
9: Geneseq20068:*
10: Geneseq20078:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	5691	100.0	5691	6	ABE63437 Human CDN
2	5332	97.2	6367	6	ABE00062 CADHP-9-c
3	793	13.9	2597	6	ABK70006 CDNA enco
4	793	13.9	2597	8	ADA01359 Human PRO
5	793	13.9	2597	8	ADA43788 Human CDN
6	793	13.9	2597	8	ADA43556 Human CDN
7	793	13.9	2597	8	ADA01231 Human PRO
8	793	13.9	2597	8	ADA01115 Human CDN
9	793	13.9	2597	8	ADA43672 Human CDN
10	793	13.9	2597	8	ADA06934 Human PRO
11	793	13.9	2597	8	ADA08422 Novel hum
12	793	13.9	2597	8	ADB99715 Human PRO
13	793	13.9	2597	8	ADB86998 Human PRO
14	793	13.9	2597	8	ADB66153 Human CDN
15	793	13.9	2597	9	ADB99831 Human PRO
16	793	13.9	2597	9	ADB99486 Novel hum
17	793	13.9	2597	9	ADB66037 Human CDN
18	793	13.9	2597	9	ADC23435 Human CDN
19	793	13.9	2597	9	ADC26128 Human PRO
20	793	13.9	2597	9	ADB04955 Human PRO
21	793	13.9	2597	9	ADB11261 Human PRO
22	793	13.9	2597	9	ADB88192 Human PRO
23	793	13.9	2597	9	ADB95487 Human CDN

24	793	13.9	2597	9	ADB06417 Human PRO
25	793	13.9	2597	9	ADB38192 Human PRO
26	793	13.9	2597	9	ADB88308 Human PRO
27	793	13.9	2597	9	ADB90889 Human CDN
28	793	13.9	2597	10	ADB51742 Human CDN
29	793	13.9	2597	10	ADB51858 Human CDN
30	793	13.9	2597	10	ADB37716 Human CDN
31	793	13.9	2597	10	ADB37600 Human CDN
32	793	13.9	2597	10	ADB95371 Human CDN
33	793	13.9	2597	10	ADB38071 Human PRO
34	793	13.9	2597	10	ADB76160 Human PRO
35	793	13.9	2597	10	ADB39483 Human PRO
36	793	13.9	2597	10	ADB04287 Human PRO
37	793	13.9	2597	10	ADB39884 Human PRO
38	793	13.9	2597	10	ADB19749 Human PRO
39	793	13.9	2597	10	ADB77327 Human CDN
40	793	13.9	2597	10	ADB65435 Human PRO
41	793	13.9	2597	10	ADB76044 Human PRO
42	793	13.9	2597	10	ADB37955 Human PRO
43	793	13.9	2597	10	ADB64565 Human PRO
44	793	13.9	2597	10	ADB38900 Human PRO
45	793	13.9	2597	10	ADB51974 Human CDN
46	793	13.9	2597	10	ADB91005 Human CDN
47	793	13.9	2597	10	ADB38784 Human PRO
48	793	13.9	2597	10	ADB37484 Human CDN
49	793	13.9	2597	10	ADB06301 Human PRO
50	793	13.9	2597	10	ADB90160 Human CDN
51	793	13.9	2597	10	ADB38668 Human PRO
52	793	13.9	2597	10	ADB39599 Human PRO
53	793	13.9	2597	10	ADB89204 Human PRO
54	793	13.9	2597	10	ADB88971 Human PRO
55	793	13.9	2597	10	ADB19865 Human PRO
56	793	13.9	2597	10	ADB77443 Human CDN
57	793	13.9	2597	10	ADB63319 Human PRO
58	793	13.9	2597	10	ADB38367 Human PRO
59	793	13.9	2597	10	ADB38552 Human CDN
60	503	8.8	598	4	AA142745 Probe #11
61	503	8.8	598	4	AAK36939 Human bon
62	503	8.8	598	4	AAK11122 Human bra
63	503	8.8	598	4	AB36611 Human liv
64	491	8.6	601	4	AA500023 Human CDN
65	481	8.5	495	4	AA155806 Probe #24
66	481	8.5	495	4	AAK49882 Human bon
67	481	8.5	495	4	AAK23792 Human bra
68	479	8.5	495	4	AB449520 Human liv
69	479	8.4	477	5	ADB62416 Human CDN
70	185	3.3	477	5	AA568294 DNA enco
71	185	3.3	477	5	ACD05873 Novel hum
72	183	3.2	477	5	AA568296 DNA enco
73	183	3.2	477	5	ACD05552 Human sec
74	183	3.2	477	5	ACD05552 Human sec
75	183	3.2	477	5	ACD05552 Human sec
76	183	3.2	477	5	ACD05552 Human sec
77	183	3.2	477	5	ACD05552 Human sec
78	183	3.2	477	5	ACD05552 Human sec
79	183	3.2	477	5	ACD05552 Human sec
80	183	3.2	477	5	ACD05552 Human sec
81	183	3.2	477	5	ACD05552 Human sec
82	183	3.2	477	5	ACD05552 Human sec
83	183	3.2	477	5	ACD05552 Human sec
84	183	3.2	477	5	ACD05552 Human sec
85	183	3.2	477	5	ACD05552 Human sec
86	183	3.2	477	5	ACD05552 Human sec
87	183	3.2	477	5	ACD05552 Human sec
88	183	3.2	477	5	ACD05552 Human sec
89	183	3.2	477	5	ACD05552 Human sec
90	183	3.2	477	5	ACD05552 Human sec
91	183	3.2	477	5	ACD05552 Human sec
92	183	3.2	477	5	ACD05552 Human sec
93	183	3.2	477	5	ACD05552 Human sec
94	183	3.2	477	5	ACD05552 Human sec
95	183	3.2	477	5	ACD05552 Human sec
96	183	3.2	477	5	ACD05552 Human sec

C 169	20	0.4	38506	7	ADA865090	S. venezu
C 168	20	0.4	38506	3	AAZ56001	Recombina
C 167	20	0.4	38506	3	AAA75633	Nucleotid
C 166	20	0.4	37948	3	AAZ87285	S. venezu
C 165	20	0.4	36778	3	AAZ87318	S. venezu
C 164	20	0.4	13844	3	AAZ87297	S. venezu
C 163	20	0.4	6561	5	AAAS85204	DNA encod
C 162	20	0.4	6329	4	AAAS85204	Human p1e
C 161	20	0.4	6329	4	AAAS85204	Human p1e
C 160	20	0.4	6252	6	AAHL67279	Thyroid c
C 159	20	0.4	6252	5	AAH81745	Human dif
C 158	20	0.4	6252	5	AAH81745	Human dif
C 157	20	0.4	3849	4	AAAF25795	S. chryso
C 156	20	0.4	2971	4	AAK93432	Human CDN
C 155	20	0.4	2564	4	ADBB63125	Human CDN
C 154	20	0.4	2522	8	ADBA48792	Novel hum
C 153	20	0.4	2522	8	ADBA48792	Novel hum
C 152	20	0.4	2444	6	AAA90083	Human neg
C 151	20	0.4	2214	6	ABK34791	Human CDN
C 150	20	0.4	2153	4	AA160597	Human ORF
C 149	20	0.4	2074	3	AACT6166	Human ORF
C 148	20	0.4	2060	4	AAH15894	Human CDN
C 147	20	0.4	1636	2	AACT9794	Mouse SH3
C 146	20	0.4	1320	4	AAAS6182	Salmonell
C 145	20	0.4	1320	4	AAAS6182	Salmonell
C 144	20	0.4	824	4	AAH05487	Human OVA
C 143	20	0.4	792	6	ABO56140	Human OVA
C 142	20	0.4	750	3	AAAO1845	Human COL
C 141	20	0.4	746	5	AAAS69769	DNA encod
C 140	20	0.4	744	3	AACT10882	Human sec
C 139	20	0.4	736	5	AAAS69768	DNA encod
C 138	20	0.4	700	6	ABN96801	Gene #129
C 137	20	0.4	689	3	AACT10893	Human sec
C 136	20	0.4	616	3	AAAO16233	Human COL
C 135	20	0.4	616	3	AAAO16233	Human COL
C 134	20	0.4	615	3	AACT00732	Human sec
C 133	20	0.4	506	9	ADPD19817	Seabass p
C 132	20	0.4	500	7	ACDP3147	Human COL
C 131	20	0.4	495	8	ACHS0755	Human nam
C 130	20	0.4	486	3	AACT07273	Human sec
C 129	20	0.4	482	8	ACH34668	Human end
C 128	20	0.4	476	8	ACH44821	Human toe
C 127	20	0.4	462	4	ABZ23809	Human COL
C 126	20	0.4	462	4	AA128623	Human COL
C 125	20	0.4	456	3	AA377885	CDNA enco
C 124	20	0.4	456	3	AACT07270	Human sec
C 123	20	0.4	412	8	ACH29208	Human adu
C 122	20	0.4	412	8	ACH29208	Human adu
C 121	20	0.4	412	3	AACT07271	Human sec
C 120	20	0.4	386	8	ACH29631	Human tes
C 119	20	0.4	382	3	AACT07272	Human sec
C 118	20	0.4	369	8	ACH21424	Human adu
C 117	20	0.4	345	5	AAAS82885	DNA encod
C 116	20	0.4	328	6	AAAS77777	CDNA #453
C 115	20	0.4	324	6	ABN94554	Gene #105
C 114	20	0.4	318	3	AAZ42602	Human 5'
C 113	20	0.4	311	7	ABX88561	CoRN ear-
C 112	21	0.4	274	6	ABL72690	CoRN taas
C 111	21	0.4	135638	9	ABX34289	S. atrocol
C 110	21	0.4	3111	9	ADBS3197	Primary r
C 109	21	0.4	3111	9	ADBS3197	Toxicity-
C 108	21	0.4	3111	9	AAAS04276	Rat gluta
C 107	21	0.4	522	5	ABV41464	Human pro
C 106	21	0.4	205	5	ABV44095	Human pro
C 105	21	0.4	205	5	ABV35258	Human pro
C 104	21	0.4	154	5	ABV04995	Human pro
C 103	21	0.4	21	6	ABBS63466	Real-time
C 102	22	0.4	22	6	ABBS63466	Real-time
C 101	23	0.4	600	6	ABSI1541	Human gen
C 100	23	0.4	600	6	ABSI1541	Human gen
C 99	23	0.4	201	6	ABBS23899	Human gen
C 98	23	0.4	201	6	ABBS23899	Human gen
C 97	23	0.4	180	3	AACT6570	Human sec

243	19	0.3	2477	8	ADB15923	Adb15923 Human PRO	316	19	0.3	2477	9	ADB74831	Adb74831 Human sec
244	19	0.3	2477	8	ADA47709	Ada47709 Human PRO	317	19	0.3	2477	9	ADB47079	Adb47079 Novel hum
245	19	0.3	2477	8	ADA18230	Ada18230 Human sec	318	19	0.3	2477	9	ADB66866	Adb66866 Human PRO
246	19	0.3	2477	8	ACD66929	Acda66929 Human CDNA	319	19	0.3	2477	9	ADB77291	Adb77291 Novel hum
247	19	0.3	2477	8	ADA67504	Ada67504 Human PRO	320	19	0.3	2477	9	ADB34448	Adb34448 Human PRO
248	19	0.3	2477	8	ADB30511	Adb30511 CDNA enco	321	19	0.3	2477	9	ADB35552	Adb35552 Human PRO
249	19	0.3	2477	8	ADA85807	Ada85807 Novel hum	322	19	0.3	2477	9	ADB33896	Adb33896 Human PRO
250	19	0.3	2477	8	ADA85807	Ada85807 Novel hum	323	19	0.3	2477	9	ADB35000	Adb35000 Human PRO
251	19	0.3	2477	8	ADA97019	Ada97019 Human PRO	324	19	0.3	2477	9	ADB36104	Adb36104 Human PRO
252	19	0.3	2477	8	ADA79323	Ada79323 Human PRO	325	19	0.3	2477	9	ADB46499	Adb46499 Novel hum
253	19	0.3	2477	8	ADA87462	Ada87462 Novel hum	326	19	0.3	2477	9	ADB28477	Adb28477 Human sec
254	19	0.3	2477	8	ADB16664	Adb16664 Human PRO	327	19	0.3	2477	9	ADC39677	Adc39677 Human sec
255	19	0.3	2477	8	ACD83090	Acda83090 Human PRO	328	19	0.3	2477	9	ADC40191	Adc40191 Human sec
256	19	0.3	2477	8	ADA16205	Ada16205 Human sec	329	19	0.3	2477	9	ADC19015	Adc19015 Human sec
257	19	0.3	2477	8	ADA91756	Ada91756 Novel hum	330	19	0.3	2477	9	ADC34315	Adc34315 Human sec
258	19	0.3	2477	8	ADB14819	Adb14819 Human PRO	331	19	0.3	2477	9	ADC29370	Adc29370 Human sec
259	19	0.3	2477	8	ADB18780	Adb18780 Novel hum	332	19	0.3	2477	9	ADC28901	Adc28901 Human sec
260	19	0.3	2477	8	ADA93995	Ada93995 Human PRO	333	19	0.3	2477	9	ADC40786	Adc40786 Human sec
261	19	0.3	2477	8	ADB19891	Adb19891 Novel hum	334	19	0.3	2477	9	ADC19443	Adc19443 Human sec
262	19	0.3	2477	8	ADB13203	Adb13203 Human PRO	335	19	0.3	2477	9	ADC38991	Adc38991 Human sec
263	19	0.3	2477	8	ACD98589	Acda98589 Novel hum	336	19	0.3	2477	9	ADC12961	Adc12961 Human sec
264	19	0.3	2477	8	ADA74457	Ada74457 Human PRO	337	19	0.3	2477	9	ADC50372	Adc50372 Novel hum
265	19	0.3	2477	8	ADA42350	Ada42350 Human sec	338	19	0.3	2477	9	ADC71919	Adc71919 Novel hum
266	19	0.3	2477	8	ADB24690	Adb24690 Human PRO	339	19	0.3	2477	9	ADC59898	Adc59898 Novel hum
267	19	0.3	2477	8	ADA82214	Ada82214 Human PRO	340	19	0.3	2477	9	ADC52905	Adc52905 Novel hum
268	19	0.3	2477	8	ADA75177	Ada75177 Human PRO	341	19	0.3	2477	9	ADC57259	Adc57259 Novel hum
269	19	0.3	2477	8	ADA85255	Ada85255 Novel hum	342	19	0.3	2477	9	ADC60450	Adc60450 Novel hum
270	19	0.3	2477	8	ADA84703	Ada84703 Novel hum	343	19	0.3	2477	9	ADC50925	Adc50925 Novel hum
271	19	0.3	2477	8	ADB29359	Adb29359 CDNA enco	344	19	0.3	2477	9	ADC65450	Adc65450 Human PRO
272	19	0.3	2477	8	ADA80487	Ada80487 Human PRO	345	19	0.3	2477	9	ADC54550	Adc54550 Novel hum
273	19	0.3	2477	8	ADA75729	Ada75729 Human PRO	346	19	0.3	2477	9	ADC53511	Adc53511 Novel hum
274	19	0.3	2477	8	ADA64954	Ada64954 Human PRO	347	19	0.3	2477	9	ADC59034	Adc59034 Novel hum
275	19	0.3	2477	8	ADB25250	Adb25250 Human PRO	348	19	0.3	2477	9	ADC58482	Adc58482 Novel hum
276	19	0.3	2477	8	ADA93426	Ada93426 Human PRO	349	19	0.3	2477	9	ADC12413	Adc12413 Human sec
277	19	0.3	2477	8	ADB26776	Adb26776 CDNA enco	350	19	0.3	2477	9	ADC03156	Adc03156 Novel hum
278	19	0.3	2477	8	ADB31063	Adb31063 CDNA enco	351	19	0.3	2477	9	ADC90148	Adc90148 Novel hum
279	19	0.3	2477	8	ADA60991	Ada60991 Homo sapi	352	19	0.3	2477	9	ADC69567	Adc69567 CDNA enco
280	19	0.3	2477	8	ADB24138	Adb24138 Human PRO	353	19	0.3	2477	9	ADC48456	Adc48456 Human PRO
281	19	0.3	2477	8	ADA96467	Ada96467 Human PRO	354	19	0.3	2477	9	ADC09985	Adc09985 Human PRO
282	19	0.3	2477	8	ADA81039	Ada81039 Human PRO	355	19	0.3	2477	9	ADC04560	Adc04560 Novel hum
283	19	0.3	2477	8	ADA95915	Ada95915 Human PRO	356	19	0.3	2477	9	ADC80516	Adc80516 Novel hum
284	19	0.3	2477	8	ADB26224	Adb26224 CDNA enco	357	19	0.3	2477	9	ADC011023	Adc011023 Human PRO
285	19	0.3	2477	8	ADB21709	Adb21709 Novel hum	358	19	0.3	2477	9	ADC47904	Adc47904 Human PRO
286	19	0.3	2477	8	ADA77488	Ada77488 Human PRO	359	19	0.3	2477	9	ADC04968	Adc04968 Human sec
287	19	0.3	2477	8	ADB18228	Adb18228 CDNA enco	360	19	0.3	2477	9	ADC79964	Adc79964 Novel hum
288	19	0.3	2477	8	ADA86911	Ada86911 Novel hum	361	19	0.3	2477	9	ADC09433	Adc09433 Human PRO
289	19	0.3	2477	8	ADA16629	Ada16629 Human sec	362	19	0.3	2477	9	ADC03974	Adc03974 Human PRO
290	19	0.3	2477	8	ADA13058	Ada13058 Human sec	363	19	0.3	2477	9	ADC03550	Adc03550 Human sec
291	19	0.3	2477	8	ADA41926	Ada41926 Human sec	364	19	0.3	2477	9	ADC54148	Adc54148 Human PRO
292	19	0.3	2477	8	ADA88014	Ada88014 Novel hum	365	19	0.3	2477	9	ADC52285	Adc52285 CDNA enco
293	19	0.3	2477	8	ADA46402	Ada46402 Novel hum	366	19	0.3	2477	9	ADC53577	Adc53577 Novel hum
294	19	0.3	2477	8	ADA17273	Ada17273 Human sec	367	19	0.3	2477	9	ADC53025	Adc53025 CDNA enco
295	19	0.3	2477	8	ADA43776	Ada43776 Human sec	368	19	0.3	2477	9	ADC51733	Adc51733 CDNA enco
296	19	0.3	2477	8	ADB28432	Adb28432 CDNA enco	369	19	0.3	2477	9	ADC02532	Adc02532 Human PRO
297	19	0.3	2477	8	ADB28984	Adb28984 CDNA enco	370	19	0.3	2477	9	ADC01196	Adc01196 Human PRO
298	19	0.3	2477	8	ADA76936	Ada76936 Human PRO	371	19	0.3	2477	9	ADC54148	Adc54148 Novel hum
299	19	0.3	2477	8	ADA88566	Ada88566 Novel hum	372	19	0.3	2477	9	ADC92465	Adc92465 Human PRO
300	19	0.3	2477	8	ADA97571	Ada97571 Human PRO	373	19	0.3	2477	9	ADC91361	Adc91361 Human PRO
301	19	0.3	2477	8	ADB27328	Adb27328 CDNA enco	374	19	0.3	2477	9	ADC03975	Adc03975 Human PRO
302	19	0.3	2477	8	ADB22261	Adb22261 Novel hum	375	19	0.3	2477	9	ADC32272	Adc32272 Novel hum
303	19	0.3	2477	8	ACD23630	Acda23630 Human PRO	376	19	0.3	2477	9	ADC22204	Adc22204 CDNA enco
304	19	0.3	2477	8	ADA66952	Ada66952 Human PRO	377	19	0.3	2477	9	ADC79428	Adc79428 Human PRO
305	19	0.3	2477	8	ADB22813	Adb22813 Human PRO	378	19	0.3	2477	9	ADC41964	Adc41964 Human PRO
306	19	0.3	2477	8	ADB23586	Adb23586 Human PRO	379	19	0.3	2477	9	ADC17781	Adc17781 Human PRO
307	19	0.3	2477	8	ADA92308	Ada92308 Novel hum	380	19	0.3	2477	9	ADC91913	Adc91913 Human PRO
308	19	0.3	2477	8	ADB15371	Adb15371 Human PRO	381	19	0.3	2477	9	ADC83376	Adc83376 Novel hum
309	19	0.3	2477	8	ADB38623	Adb38623 Novel hum	382	19	0.3	2477	9	ADC33928	Adc33928 Novel hum
310	19	0.3	2477	8	ADB38071	Adb38071 Novel hum	383	19	0.3	2477	9	ADC79980	Adc79980 CDNA enco
311	19	0.3	2477	8	ADB66543	Adb66543 Novel hum	384	19	0.3	2477	9	ADC93017	Adc93017 Human PRO
312	19	0.3	2477	8	ADB89623	Adb89623 Human PRO	385	19	0.3	2477	9	ADC19437	Adc19437 Human PRO
313	19	0.3	2477	8	ADB90355	Adb90355 Human PRO	386	19	0.3	2477	9	ADC34802	Adc34802 Human sec
314	19	0.3	2477	8	ADB77695	Adb77695 Human sec	387	19	0.3	2477	9	ADC18885	Adc18885 Human PRO
315	19	0.3	2477	9	ADB39456	Adb39456 Novel hum	388	19	0.3	2477	9	ADB74831	Adb74831 Human sec

389	19	0.3	2477	9	ADBA33081	462	18	0.3	164	4	ABAA37215	AbAa37215
390	19	0.3	2477	9	ADBA35870	463	18	0.3	164	4	AAK44864	AaK44864
391	19	0.3	2477	9	ADBA22756	464	18	0.3	164	4	AAK18919	AaK18919
392	19	0.3	2477	9	ADBD78874	465	18	0.3	164	4	ABSA45526	AbSa45526
393	19	0.3	2477	9	ADBA28284	466	18	0.3	164	6	ABSA19105	AbSa19105
394	19	0.3	2477	9	ADBA25156	467	18	0.3	209	4	ABDL22363	AbDL22363
395	19	0.3	2477	9	ADBA05532	468	18	0.3	248	4	AAAL00188	AAAL00188
396	19	0.3	2477	9	ADBA05532	469	18	0.3	258	7	ABXA91685	AbXa91685
397	19	0.3	2477	9	ADBA05532	470	18	0.3	262	7	AAAA42325	AaAa42325
398	19	0.3	2477	9	ADBA05532	471	18	0.3	278	2	AAAT22474	AAAT22474
399	19	0.3	2477	10	ADBA05532	472	18	0.3	288	7	ABXA87821	AbXa87821
400	19	0.3	2477	10	ADBA05532	473	18	0.3	298	4	AAFA31739	AAFA31739
401	19	0.3	2477	10	ADBA05532	474	18	0.3	326	4	AAHS08812	AAHS08812
402	19	0.3	2477	10	ADBA05532	475	18	0.3	326	4	AAHS08810	AAHS08810
403	19	0.3	2477	10	ADBA05532	476	18	0.3	382	6	ABKA64256	ABKA64256
404	19	0.3	2477	10	ADBA05532	477	18	0.3	384	6	ABN21187	ABN21187
405	19	0.3	2477	10	ADBA05532	478	18	0.3	414	6	ABKA5931	ABKA5931
406	19	0.3	2477	10	ADBA05532	479	18	0.3	420	6	ABKA5931	ABKA5931
407	19	0.3	2477	10	ADBA05532	480	18	0.3	429	4	ABAA5948	ABAA5948
408	19	0.3	2477	10	ADBA05532	481	18	0.3	429	4	AAI33602	AAI33602
409	19	0.3	2477	10	ADBA05532	482	18	0.3	429	4	ABAA44507	ABAA44507
410	19	0.3	2477	10	ADBA05532	483	18	0.3	429	4	ABAA24714	ABAA24714
411	19	0.3	2477	10	ADBA05532	484	18	0.3	429	4	AAK28673	AAK28673
412	19	0.3	2477	10	ADBA05532	485	18	0.3	429	4	AAK03220	AAK03220
413	19	0.3	2477	10	ADBA05532	486	18	0.3	429	4	AAK03220	AAK03220
414	19	0.3	2477	10	ADBA05532	487	18	0.3	429	4	AAK03220	AAK03220
415	19	0.3	2477	10	ADBA05532	488	18	0.3	429	4	AAK03220	AAK03220
416	19	0.3	2477	10	ADBA05532	489	18	0.3	429	4	AAK03220	AAK03220
417	19	0.3	2477	10	ADBA05532	490	18	0.3	429	4	AAK03220	AAK03220
418	19	0.3	2477	10	ADBA05532	491	18	0.3	429	4	AAK03220	AAK03220
419	19	0.3	2477	10	ADBA05532	492	18	0.3	429	4	AAK03220	AAK03220
420	19	0.3	2477	10	ADBA05532	493	18	0.3	429	4	AAK03220	AAK03220
421	19	0.3	2477	10	ADBA05532	494	18	0.3	429	4	AAK03220	AAK03220
422	19	0.3	2477	10	ADBA05532	495	18	0.3	429	4	AAK03220	AAK03220
423	19	0.3	2477	10	ADBA05532	496	18	0.3	429	4	AAK03220	AAK03220
424	19	0.3	2477	10	ADBA05532	497	18	0.3	429	4	AAK03220	AAK03220
425	19	0.3	2477	10	ADBA05532	498	18	0.3	429	4	AAK03220	AAK03220
426	19	0.3	2477	10	ADBA05532	499	18	0.3	429	4	AAK03220	AAK03220
427	19	0.3	2477	10	ADBA05532	500	18	0.3	429	4	AAK03220	AAK03220
428	19	0.3	2477	10	ADBA05532	501	18	0.3	429	4	AAK03220	AAK03220
429	19	0.3	2477	10	ADBA05532	502	18	0.3	429	4	AAK03220	AAK03220
430	19	0.3	2477	10	ADBA05532	503	18	0.3	429	4	AAK03220	AAK03220
431	19	0.3	2477	10	ADBA05532	504	18	0.3	429	4	AAK03220	AAK03220
432	19	0.3	2477	10	ADBA05532	505	18	0.3	429	4	AAK03220	AAK03220
433	19	0.3	2477	10	ADBA05532	506	18	0.3	429	4	AAK03220	AAK03220
434	19	0.3	2477	10	ADBA05532	507	18	0.3	429	4	AAK03220	AAK03220
435	19	0.3	2477	10	ADBA05532	508	18	0.3	429	4	AAK03220	AAK03220
436	19	0.3	2477	10	ADBA05532	509	18	0.3	429	4	AAK03220	AAK03220
437	19	0.3	2477	10	ADBA05532	510	18	0.3	429	4	AAK03220	AAK03220
438	19	0.3	2477	10	ADBA05532	511	18	0.3	429	4	AAK03220	AAK03220
439	19	0.3	2477	10	ADBA05532	512	18	0.3	429	4	AAK03220	AAK03220
440	19	0.3	2477	10	ADBA05532	513	18	0.3	429	4	AAK03220	AAK03220
441	19	0.3	2477	10	ADBA05532	514	18	0.3	429	4	AAK03220	AAK03220
442	19	0.3	2477	10	ADBA05532	515	18	0.3	429	4	AAK03220	AAK03220
443	19	0.3	2477	10	ADBA05532	516	18	0.3	429	4	AAK03220	AAK03220
444	19	0.3	2477	10	ADBA05532	517	18	0.3	429	4	AAK03220	AAK03220
445	19	0.3	2477	10	ADBA05532	518	18	0.3	429	4	AAK03220	AAK03220
446	19	0.3	2477	10	ADBA05532	519	18	0.3	429	4	AAK03220	AAK03220
447	19	0.3	2477	10	ADBA05532	520	18	0.3	429	4	AAK03220	AAK03220
448	19	0.3	2477	10	ADBA05532	521	18	0.3	429	4	AAK03220	AAK03220
449	19	0.3	2477	10	ADBA05532	522	18	0.3	429	4	AAK03220	AAK03220
450	19	0.3	2477	10	ADBA05532	523	18	0.3	429	4	AAK03220	AAK03220
451	19	0.3	2477	10	ADBA05532	524	18	0.3	429	4	AAK03220	AAK03220
452	19	0.3	2477	10	ADBA05532	525	18	0.3	429	4	AAK03220	AAK03220
453	19	0.3	2477	10	ADBA05532	526	18	0.3	429	4	AAK03220	AAK03220
454	19	0.3	2477	10	ADBA05532	527	18	0.3	429	4	AAK03220	AAK03220
455	19	0.3	2477	10	ADBA05532	528	18	0.3	429	4	AAK03220	AAK03220
456	19	0.3	2477	10	ADBA05532	529	18	0.3	429	4	AAK03220	AAK03220
457	19	0.3	2477	10	ADBA05532	530	18	0.3	429	4	AAK03220	AAK03220
458	19	0.3	2477	10	ADBA05532	531	18	0.3	429	4	AAK03220	AAK03220
459	19	0.3	2477	10	ADBA05532	532	18	0.3	429	4	AAK03220	AAK03220
460	19	0.3	2477	10	ADBA05532	533	18	0.3	429	4	AAK03220	AAK03220
461	19	0.3	2477	10	ADBA05532	534	18	0.3	429	4	AAK03220	AAK03220

C 535	18	0.3	729	8	ADB08473	ADB08473	Alloioccc	C 608	18	0.3	961	9	ADC17278	ADC17278	cdna sequ
C 536	18	0.3	742	9	ADB67959	ADB67959	Human lun	C 609	18	0.3	961	9	ADC14976	ADC14976	Novel hum
C 537	18	0.3	746	2	AAx40036	AAx40036	Prostate	C 610	18	0.3	961	9	ADC34950	ADC34950	Human PRO
C 538	18	0.3	759	6	ABK30714	ABK30714	Plant dwa	C 611	18	0.3	961	9	ADC52471	ADC52471	Novel hum
C 539	18	0.3	759	6	AB149515	AB149515	Sequence	C 612	18	0.3	961	9	ADC21940	ADC21940	Human PRO
C 540	18	0.3	797	9	ADBS6800	ADBS6800	Toxicity-	C 613	18	0.3	961	9	ADC49971	ADC49971	Novel hum
C 541	18	0.3	797	9	ADBS3098	ADBS3098	Primary r	C 614	18	0.3	961	9	ADC49170	ADC49170	Novel hum
C 542	18	0.3	808	4	AAH06598	AAH06598	Human CDN	C 615	18	0.3	961	9	ADC49687	ADC49687	Novel hum
C 543	18	0.3	814	5	AAH11362	AAH11362	Soybean D	C 616	18	0.3	961	9	ADC47548	ADC47548	Novel hum
C 544	18	0.3	816	4	ABA69602	ABA69602	Human toe	C 617	18	0.3	961	9	ADC47293	ADC47293	Novel hum
C 545	18	0.3	816	4	ABSA4312	ABSA4312	Human liv	C 618	18	0.3	961	9	ADC78168	ADC78168	Novel hum
C 546	18	0.3	831	6	ABDL17516	ABDL17516	Segt	C 619	18	0.3	961	9	ADC06403	ADC06403	Novel hum
C 547	18	0.3	832	6	ABK09490	ABK09490	Human ova	C 620	18	0.3	961	9	ADC77922	ADC77922	Novel hum
C 548	18	0.3	833	6	AAx40033	AAx40033	Prostate	C 621	18	0.3	961	9	ADD50885	ADD50885	Novel hum
C 549	18	0.3	840	6	ABQ68625	ABQ68625	Listeria	C 622	18	0.3	961	9	ADD51131	ADD51131	Novel hum
C 550	18	0.3	845	4	AAH98771	AAH98771	Human, Bst	C 623	18	0.3	961	9	ADD50612	ADD50612	Novel hum
C 551	18	0.3	849	6	ABK75468	ABK75468	Bacillus	C 624	18	0.3	961	9	ADD50366	ADD50366	Novel hum
C 552	18	0.3	873	6	ABSG67310	ABSG67310	Breast sp	C 625	18	0.3	961	9	ADD51377	ADD51377	Novel hum
C 553	18	0.3	882	5	AAD08650	AAD08650	Mouse can	C 626	18	0.3	961	9	ADD52281	ADD52281	Novel hum
C 554	18	0.3	899	7	ABT15733	ABT15733	Human can	C 627	18	0.3	961	10	ADD52895	ADD52895	Novel hum
C 555	18	0.3	921	5	AAAS87405	AAAS87405	Human sec	C 628	18	0.3	961	10	ADD50939	ADD50939	Human PRO
C 556	18	0.3	921	5	AAAS87191	AAAS87191	Gene enco	C 629	18	0.3	961	10	ADD75168	ADD75168	Human PRO
C 557	18	0.3	921	7	ADAS5984	ADAS5984	Human sec	C 630	18	0.3	961	10	ADD75914	ADD75914	Novel hum
C 558	18	0.3	921	7	ADAS39795	ADAS39795	Human sec	C 631	18	0.3	961	10	ADD85146	ADD85146	Novel hum
C 559	18	0.3	921	7	ACCS0432	ACCS0432	Human sec	C 632	18	0.3	961	10	ADD86972	ADD86972	Novel hum
C 560	18	0.3	921	7	ADCF7491	ADCF7491	Human sec	C 633	18	0.3	961	10	ADD86972	ADD86972	Novel hum
C 561	18	0.3	921	7	ADCF7491	ADCF7491	Human sec	C 634	18	0.3	961	10	ADD86972	ADD86972	Novel hum
C 562	18	0.3	961	4	AAFP92129	AAFP92129	Human PRO	C 635	18	0.3	961	10	ADBE20849	ADBE20849	Novel hum
C 563	18	0.3	961	6	ABK33645	ABK33645	CDNA enco	C 636	18	0.3	961	10	ADBE39146	ADBE39146	Novel hum
C 564	18	0.3	961	6	ABST74449	ABST74449	Human CDN	C 637	18	0.3	961	10	ADBE05693	ADBE05693	Human PRO
C 565	18	0.3	961	7	ACAE6951	ACAE6951	CDNA enco	C 638	18	0.3	961	10	ADD73678	ADD73678	Human PRO
C 566	18	0.3	961	7	ACAE6703	ACAE6703	Novel hum	C 639	18	0.3	961	10	ADD78518	ADD78518	Novel hum
C 567	18	0.3	961	7	ACAE91235	ACAE91235	Novel hum	C 640	18	0.3	961	10	ADDE21341	ADDE21341	Novel hum
C 568	18	0.3	961	7	ACAE81612	ACAE81612	Human CDN	C 641	18	0.3	961	10	ADDE21341	ADDE21341	Novel hum
C 569	18	0.3	961	7	ACAE60434	ACAE60434	Novel hum	C 642	18	0.3	961	10	ADDE20603	ADDE20603	Novel hum
C 570	18	0.3	961	7	ACAE58881	ACAE58881	CDNA enco	C 643	18	0.3	961	10	ADD75668	ADD75668	Human PRO
C 571	18	0.3	961	7	ACAE6057	ACAE6057	CDNA enco	C 644	18	0.3	961	10	ADD74184	ADD74184	Human PRO
C 572	18	0.3	961	7	ACAE91321	ACAE91321	CDNA enco	C 645	18	0.3	961	10	ADD74430	ADD74430	Human PRO
C 573	18	0.3	961	7	ACAE45220	ACAE45220	Human sec	C 646	18	0.3	961	10	ADD75160	ADD75160	Novel hum
C 574	18	0.3	961	7	ACAE93768	ACAE93768	Human CDN	C 647	18	0.3	961	10	ADD86562	ADD86562	Novel hum
C 575	18	0.3	961	7	ACAE67342	ACAE67342	CDNA enco	C 648	18	0.3	961	10	ADDE05201	ADDE05201	Human PRO
C 576	18	0.3	961	7	ACAE68607	ACAE68607	Novel hum	C 649	18	0.3	961	10	ADD759414	ADD759414	Human PRO
C 577	18	0.3	961	7	ACH66315	ACH66315	Novel hum	C 650	18	0.3	961	10	ADD76958	ADD76958	Novel hum
C 578	18	0.3	961	7	ACHD03369	ACHD03369	Novel hum	C 651	18	0.3	961	10	ADD86726	ADD86726	Novel hum
C 579	18	0.3	961	7	ACHA89360	ACHA89360	Novel hum	C 652	18	0.3	961	10	ADD78194	ADD78194	Novel hum
C 580	18	0.3	961	7	ACHA68997	ACHA68997	Novel hum	C 653	18	0.3	961	10	ADD77702	ADD77702	Novel hum
C 581	18	0.3	961	7	ACHA98519	ACHA98519	Human PRO	C 654	18	0.3	961	10	ADD77948	ADD77948	Novel hum
C 582	18	0.3	961	8	ACHA63444	ACHA63444	CDNA enco	C 655	18	0.3	961	10	ADD85406	ADD85406	Novel hum
C 583	18	0.3	961	8	ABT44336	ABT44336	Human PRO	C 656	18	0.3	961	10	ADD73938	ADD73938	Novel hum
C 584	18	0.3	961	8	ABD17200	ABD17200	Human CDN	C 657	18	0.3	961	10	ADD74676	ADD74676	Novel hum
C 585	18	0.3	961	8	ACH03647	ACH03647	Human sec	C 658	18	0.3	961	10	ADD77204	ADD77204	Novel hum
C 586	18	0.3	961	8	ADAE20005	ADAE20005	Novel hum	C 659	18	0.3	961	10	ADD85898	ADD85898	Novel hum
C 587	18	0.3	961	8	ADBI17388	ADBI17388	Human CDN	C 660	18	0.3	961	10	ADDE05447	ADDE05447	Novel hum
C 588	18	0.3	961	8	ADCD82161	ADCD82161	Human sec	C 661	18	0.3	961	10	ADD74922	ADD74922	Novel hum
C 589	18	0.3	961	8	ADT44619	ADT44619	Human PRO	C 662	18	0.3	961	10	ADD74922	ADD74922	Novel hum
C 590	18	0.3	961	8	ADCD82286	ADCD82286	Human sec	C 663	18	0.3	961	10	ADD74922	ADD74922	Novel hum
C 591	18	0.3	961	8	ADA00474	ADA00474	Human sec	C 664	18	0.3	961	10	ADD74922	ADD74922	Novel hum
C 592	18	0.3	961	8	ABT43992	ABT43992	Human mem	C 665	18	0.3	961	10	ADD74922	ADD74922	Novel hum
C 593	18	0.3	961	8	ADBS8716	ADBS8716	Novel hum	C 666	18	0.3	961	10	ADD74922	ADD74922	Novel hum
C 594	18	0.3	961	8	ADBS8711	ADBS8711	Novel hum	C 667	18	0.3	961	10	ADD74922	ADD74922	Novel hum
C 595	18	0.3	961	8	ADBS80817	ADBS80817	Novel hum	C 668	18	0.3	961	10	ADD74922	ADD74922	Novel hum
C 596	18	0.3	961	8	ADBS73358	ADBS73358	Novel hum	C 669	18	0.3	961	10	ADD74922	ADD74922	Novel hum
C 597	18	0.3	961	8	ADBS78440	ADBS78440	Novel hum	C 670	18	0.3	961	10	ADD74922	ADD74922	Novel hum
C 598	18	0.3	961	8	ADBS5088	ADBS5088	Human PRO	C 671	18	0.3	961	10	ADD74922	ADD74922	Novel hum
C 599	18	0.3	961	8	ADBS87194	ADBS87194	Novel hum	C 672	18	0.3	961	10	ADD74922	ADD74922	Novel hum
C 600	18	0.3	961	8	ADBS87260	ADBS87260	Human PRO	C 673	18	0.3	961	10	ADD74922	ADD74922	Novel hum
C 601	18	0.3	961	8	ADBS84642	ADBS84642	Human PRO	C 674	18	0.3	961	10	ADD74922	ADD74922	Novel hum
C 602	18	0.3	961	8	ADBS68395	ADBS68395	Human PRO	C 675	18	0.3	961	10	ADD74922	ADD74922	Novel hum
C 603	18	0.3	961	8	ADBS68202	ADBS68202	Human PRO	C 676	18	0.3	961	10	ADD74922	ADD74922	Novel hum
C 604	18	0.3	961	8	ADBS83957	ADBS83957	Novel hum	C 677	18	0.3	961	10	ADD74922	ADD74922	Novel hum
C 605	18	0.3	961	8	ADBS73112	ADBS73112	Novel hum	C 678	18	0.3	961	10	ADD74922	ADD74922	Novel hum
C 606	18	0.3	961	8	ADBS91019	ADBS91019	Novel hum	C 679	18	0.3	961	10	ADD74922	ADD74922	Novel hum
C 607	18	0.3	961	8	ADBS07099	ADBS07099	Human PRO	C 680	18	0.3	961	10	ADD74922	ADD74922	Novel hum

C 681	18	0.3	1236	4	AAK78770	Aak78770	Human	imm	C 754	18	0.3	2724	4	ABLI6706	ABLI6706	Drosophila
C 682	18	0.3	1236	4	AAK78769	Aak78769	Human	imm	C 755	18	0.3	2810	4	ABZ24708	ABZ24708	Human cel
C 683	18	0.3	1242	5	AAK87318	AAK87318	DNA	encod	C 756	18	0.3	2888	7	ABLI6825	ABLI6825	Drosophila
C 684	18	0.3	1242	6	AAK93932	AAK93932	DNA of th		C 757	18	0.3	2940	2	AAK60397	AAK60397	Human pol
C 685	18	0.3	1317	7	ACA36042	ACA36042	Prokaryot		C 758	18	0.3	3000	2	AAK84399	AAK84399	S. caperu
C 686	18	0.3	1317	7	ACA31646	ACA31646	Prokaryot		C 759	18	0.3	3044	4	ABLI06599	ABLI06599	Drosophila
C 687	18	0.3	1335	4	AAK61707	AAK61707	Human imm		C 760	18	0.3	3093	6	ABZ13765	ABZ13765	Arabidops
C 688	18	0.3	1362	7	ACA30833	ACA30833	Prokaryot		C 761	18	0.3	3183	7	AAI52258	AAI52258	184PIE2-r
C 689	18	0.3	1385	4	AAH99262	AAH99262	Human pro		C 762	18	0.3	3183	7	AAI52259	AAI52259	184PIE2 g
C 690	18	0.3	1395	8	ABE12712	ABE12712	Alloioococ		C 763	18	0.3	3183	7	AAI52262	AAI52262	184PIE2-r
C 691	18	0.3	1405	6	AAI48815	AAI48815	Murine an		C 764	18	0.3	3183	7	AAI52268	AAI52268	184PIE2-r
C 692	18	0.3	1406	7	ABX05078	ABX05078	Human nov		C 765	18	0.3	3183	7	AAI52260	AAI52260	184PIE2 g
C 693	18	0.3	1425	8	ADA03012	ADA03012	Mouse Ncf		C 766	18	0.3	3183	7	AAI52265	AAI52265	184PIE2 g
C 694	18	0.3	1425	9	ADB72750	ADB72750	Mouse Ncf		C 767	18	0.3	3183	7	AAI52263	AAI52263	184PIE2 g
C 695	18	0.3	1425	9	ADB72750	ADB72750	Mouse Ncf		C 768	18	0.3	3183	7	AAI52267	AAI52267	184PIE2 g
C 696	18	0.3	1425	9	ADB72750	ADB72750	Mouse Ncf		C 769	18	0.3	3183	7	AAI52264	AAI52264	184PIE2 g
C 697	18	0.3	1427	7	ACD07381	ACD07381	Rice Myb-		C 770	18	0.3	3183	7	AAI52261	AAI52261	184PIE2 g
C 698	18	0.3	1460	2	AAI38075	AAI38075	Sequence		C 771	18	0.3	3183	7	AAI52266	AAI52266	184PIE2 g
C 699	18	0.3	1461	8	ACD26032	ACD26032	Human nov		C 772	18	0.3	3227	4	AAK80792	AAK80792	Human imm
C 700	18	0.3	1507	8	AD30526	AD30526	Human nov		C 773	18	0.3	3227	4	AAK80791	AAK80791	Human imm
C 701	18	0.3	1517	6	ABO70450	ABO70450	Wild-type		C 774	18	0.3	3227	4	AAK80791	AAK80791	Human imm
C 702	18	0.3	1520	3	AB52341	AB52341	Fruit fly		C 775	18	0.3	3251	4	ABLI13875	ABLI13875	Drosophila
C 703	18	0.3	1520	6	AB53441	AB53441	Fruit fly		C 776	18	0.3	3347	4	ABLI19736	ABLI19736	Drosophila
C 704	18	0.3	1554	7	ACF70169	ACF70169	Phototrab		C 777	18	0.3	3347	4	ABLI19736	ABLI19736	Drosophila
C 705	18	0.3	1559	4	ABLI23662	ABLI23662	Drosophila		C 778	18	0.3	3366	6	ABE07269	ABE07269	Drosophila
C 706	18	0.3	1581	5	AAH21791	AAH21791	Mouse G4S		C 779	18	0.3	3408	9	AAK80791	AAK80791	Human imm
C 707	18	0.3	1604	9	ADB59144	ADB59144	Toxicity-		C 780	18	0.3	3437	2	AAZ07545	AAZ07545	Human vit
C 708	18	0.3	1607	9	ADB59144	ADB59144	Toxicity-		C 781	18	0.3	3510	9	AAZ07545	AAZ07545	Human vit
C 709	18	0.3	1609	3	ACG53307	ACG53307	Arabidops		C 782	18	0.3	3622	1	AAI60204	AAI60204	Drosophila
C 710	18	0.3	1617	4	AAI12576	AAI12576	Human pro		C 783	18	0.3	3659	4	AAI60204	AAI60204	Drosophila
C 711	18	0.3	1633	7	ABQ77395	ABQ77395	Human RNA		C 784	18	0.3	3688	4	ADB72946	ADB72946	Human kin
C 712	18	0.3	1648	6	ABQ60809	ABQ60809	Human RNA		C 785	18	0.3	3854	6	AAH17848	AAH17848	Human CDN
C 713	18	0.3	1751	4	AAI58611	AAI58611	Human pol		C 786	18	0.3	3969	6	AAH17848	AAH17848	Human nuc
C 714	18	0.3	1751	8	ADB48589	ADB48589	Novel hum		C 787	18	0.3	4077	8	ACF35658	ACF35658	Human kin
C 715	18	0.3	1771	8	ADB48589	ADB48589	Toxicity-		C 788	18	0.3	4189	7	AAZ07545	AAZ07545	Human vit
C 716	18	0.3	1771	8	ADB48589	ADB48589	Toxicity-		C 789	18	0.3	4202	7	ABLI17590	ABLI17590	Drosophila
C 717	18	0.3	1771	8	ADB48589	ADB48589	Toxicity-		C 790	18	0.3	4233	4	AAK53303	AAK53303	Human pol
C 718	18	0.3	1771	8	ADB48589	ADB48589	Toxicity-		C 791	18	0.3	4287	4	AAK53303	AAK53303	Human pol
C 719	18	0.3	1827	3	AAFI18251	AAFI18251	Lung cano		C 792	18	0.3	4304	4	AAK53303	AAK53303	Human pol
C 720	18	0.3	1840	3	AAK94394	AAK94394	Human ful		C 793	18	0.3	4408	7	AAI02765	AAI02765	Human vit
C 721	18	0.3	1882	2	AAI08702	AAI08702	Rat OCT-1		C 794	18	0.3	4461	9	ADB99955	ADB99955	Human vit
C 722	18	0.3	1882	2	AAI08702	AAI08702	Rat OCT-1		C 795	18	0.3	4461	9	ADB99955	ADB99955	Human vit
C 723	18	0.3	1882	2	AAI08702	AAI08702	Rat OCT-1		C 796	18	0.3	4494	9	ADB99955	ADB99955	Human vit
C 724	18	0.3	1882	2	AAI08702	AAI08702	Rat OCT-1		C 797	18	0.3	4536	2	AAV03326	AAV03326	DNA encod
C 725	18	0.3	1926	2	AAV82522	AAV82522	Primary r		C 798	18	0.3	4596	5	AAV03326	AAV03326	DNA encod
C 726	18	0.3	1948	4	ABAS7042	ABAS7042	Human lve		C 799	18	0.3	4596	5	AAV03326	AAV03326	DNA encod
C 727	18	0.3	1948	4	ABAS7042	ABAS7042	Human lve		C 800	18	0.3	4596	5	AAV03326	AAV03326	DNA encod
C 728	18	0.3	1959	5	AAI26317	AAI26317	Pseudomon		C 801	18	0.3	4596	5	AAV03326	AAV03326	DNA encod
C 729	18	0.3	2000	4	ABLI20725	ABLI20725	Drosophila		C 802	18	0.3	4596	5	AAV03326	AAV03326	DNA encod
C 730	18	0.3	2006	9	ADC3458	ADC3458	Human nov		C 803	18	0.3	4604	5	AAV41327	AAV41327	Human vit
C 731	18	0.3	2014	3	ACA23442	ACA23442	Human sec		C 804	18	0.3	4604	5	AAV41327	AAV41327	Human vit
C 732	18	0.3	2084	7	ABLI03833	ABLI03833	Drosophila		C 805	18	0.3	4604	5	AAV41327	AAV41327	Human vit
C 733	18	0.3	2213	4	ABLI03833	ABLI03833	Drosophila		C 806	18	0.3	4604	5	AAV41327	AAV41327	Human vit
C 734	18	0.3	2213	4	ABLI03833	ABLI03833	Drosophila		C 807	18	0.3	4604	5	AAV41327	AAV41327	Human vit
C 735	18	0.3	2253	6	ABZ14355	ABZ14355	Arabidops		C 808	18	0.3	4673	2	AAI12787	AAI12787	Pea phyto
C 736	18	0.3	2321	4	ABLI20727	ABLI20727	Drosophila		C 809	18	0.3	4727	4	ABLI13386	ABLI13386	Human vit
C 737	18	0.3	2335	7	ACCE51207	ACCE51207	Human pik		C 810	18	0.3	4794	4	ABLI19928	ABLI19928	Human vit
C 738	18	0.3	2335	7	ACCE51207	ACCE51207	Human pik		C 811	18	0.3	4794	4	ABLI19928	ABLI19928	Human vit
C 739	18	0.3	2346	4	ABLI20130	ABLI20130	Drosophila		C 812	18	0.3	4847	2	AAV58943	AAV58943	Drosophila
C 740	18	0.3	2379	6	ABZ14366	ABZ14366	Arabidops		C 813	18	0.3	5019	4	AAI04636	AAI04636	Human tes
C 741	18	0.3	2414	6	ABLI22150	ABLI22150	Drosophila		C 814	18	0.3	5049	4	ABLI97603	ABLI97603	Human tes
C 742	18	0.3	2415	6	ABK73868	ABK73868	Bacillus		C 815	18	0.3	5049	4	ABLI97603	ABLI97603	Human tes
C 743	18	0.3	2430	5	AA590898	AA590898	DNA encod		C 816	18	0.3	5135	4	ABLI03464	ABLI03464	Drosophila
C 744	18	0.3	2484	4	ABLI03465	ABLI03465	Drosophila		C 817	18	0.3	5168	6	ABK94716	ABK94716	Human tes
C 745	18	0.3	2489	4	ABLI22152	ABLI22152	Drosophila		C 818	18	0.3	5211	5	AA565560	AA565560	DNA encod
C 746	18	0.3	2505	2	AAV69965	AAV69965	Human gro		C 819	18	0.3	5289	9	ADB69194	ADB69194	C. neofo
C 747	18	0.3	2510	6	ABK83456	ABK83456	Human CDN		C 820	18	0.3	5350	4	ABLI06599	ABLI06599	Drosophila
C 748	18	0.3	2544	9	ADC30652	ADC30652	Human nov		C 821	18	0.3	5350	4	ABLI06599	ABLI06599	Drosophila
C 749	18	0.3	2596	4	AAI186538	AAI186538	Human pol		C 822	18	0.3	5353	9	ADB53664	ADB53664	Human pro
C 750	18	0.3	2602	4	ABLI26431	ABLI26431	Drosophila		C 823	18	0.3	5433	9	ADB53664	ADB53664	Human pro
C 751	18	0.3	2613	6	ABK28647	ABK28647	Human CDN		C 824	18	0.3	5484	4	ABLI13874	ABLI13874	Human SH2
C 752	18	0.3	2626	6	ABK28647	ABK28647	Human CDN		C 825	18	0.3	5557	5	AAH21794	AAH21794	Human SH2
C 753	18	0.3	2670	9	ADB69916	ADB69916	C. neofo		C 826	18	0.3	5699	5	AA589500	AA589500	DNA encod

827	18	0.3	5721	4	ABLI1973	AbI11973 Drosophila	900	18	0.3	110000	6	ABA03041_03	Continuation (4 of
C 828	18	0.3	5938	7	ABK84688	Abt19922 Apegerg11	C 901	18	0.3	110000	7	ACF65385_3	Continuation (4 of
C 829	18	0.3	6186	7	ABT19929	Abt19922 Apegerg11	C 902	18	0.3	110000	7	ACF67367_32	Continuation (33 of
C 830	18	0.3	6186	7	ABT19929	Abt19922 Apegerg11	C 903	18	0.3	110000	7	ABD12064_06	Continuation (7 of
C 831	18	0.3	6242	9	ADD71201	Human int	C 904	18	0.3	110000	8	ABD12064_15	Continuation (16 of
C 832	18	0.3	6601	7	ABT19996	Abt19996 Apegerg11	C 905	18	0.3	122186	4	AAH88703	Human his
C 833	18	0.3	6601	7	ABT19996	Abt19996 Apegerg11	C 906	18	0.3	122186	4	AAH88703	Human DNA
C 834	18	0.3	6601	7	ABT19996	Abt19996 Apegerg11	C 907	18	0.3	165199	6	ABK83460	Human CDN
C 835	18	0.3	6601	7	ABT19996	Abt19996 Apegerg11	C 908	18	0.3	188888	6	ABQ75562	Human rel
C 836	18	0.3	6690	4	ABLI20726	AbI16769 Drosophila	C 909	18	0.3	197997	7	AAI54074	Human tra
C 837	18	0.3	7033	3	AAZ36989	AbI16769 Drosophila	C 910	18	0.3	198161	6	ABK83564	Human CDN
C 838	18	0.3	7060	4	ABLI20724	AbI16769 Drosophila	C 911	18	0.3	220895	6	ABK84798	Human CDN
C 839	18	0.3	7079	4	ABLI20724	AbI16769 Drosophila	C 912	18	0.3	349980	6	ABQ81847	Human his
C 840	18	0.3	7228	4	ABLI20724	AbI16769 Drosophila	C 913	18	0.3	349980	6	ABQ81847	Human DNA
C 841	18	0.3	7330	4	ABLI20724	AbI16769 Drosophila	C 914	18	0.3	349980	6	ABQ81847	Human DNA
C 842	18	0.3	7568	5	AAH80823	AAH80823 DNA encod	C 915	18	0.3	349980	6	ABQ81847	Human DNA
C 843	18	0.3	8087	4	ABLI1972	AbI1972 Drosophila	C 916	18	0.3	349980	6	ABQ81847	Human DNA
C 844	18	0.3	8254	5	ABAI17475	AbAI17475 Human ner	C 917	18	0.3	349980	6	ABQ81847	Human DNA
C 845	18	0.3	9220	5	ABV24414	AbV24414 Human pro	C 918	18	0.3	349980	6	ABQ81847	Human DNA
C 846	18	0.3	9220	5	ABV25265	AbV25265 Human pro	C 919	18	0.3	349980	6	ABQ81847	Human DNA
C 847	18	0.3	9416	6	ABK83800	AbK83800 Human CDN	C 920	18	0.3	349980	6	ABQ81847	Human DNA
C 848	18	0.3	9416	6	ABK83800	AbK83800 Human CDN	C 921	18	0.3	349980	6	ABQ81847	Human DNA
C 849	18	0.3	10543	7	ABLC6498	AbI106498 Drosophila	C 922	18	0.3	349980	6	ABQ81847	Human DNA
C 850	18	0.3	11304	4	AAI03113	AAI03113 Human rep	C 923	18	0.3	349980	6	ABQ81847	Human DNA
C 851	18	0.3	11950	9	ADBS3871	ADBS3871 Human pro	C 924	18	0.3	349980	6	ABQ81847	Human DNA
C 852	18	0.3	13993	9	ADBS31304	ADBS31304 Testeter	C 925	18	0.3	349980	6	ABQ81847	Human DNA
C 853	18	0.3	14070	2	AAZ09525	AAZ09525 Human Apo	C 926	18	0.3	349980	6	ABQ81847	Human DNA
C 854	18	0.3	14070	2	AAZ09525	AAZ09525 Human Apo	C 927	18	0.3	349980	6	ABQ81847	Human DNA
C 855	18	0.3	14121	6	ABT13015	Abt13015 Human apo	C 928	18	0.3	349980	6	ABQ81847	Human DNA
C 856	18	0.3	14121	6	ABT13015	Abt13015 Human apo	C 929	18	0.3	349980	6	ABQ81847	Human DNA
C 857	18	0.3	15254	4	AAK83135	AAK83135 Human imm	C 930	18	0.3	349980	6	ABQ81847	Human DNA
C 858	18	0.3	15254	4	AAK83135	AAK83135 Human imm	C 931	18	0.3	349980	6	ABQ81847	Human DNA
C 859	18	0.3	15254	4	AAK83135	AAK83135 Human imm	C 932	18	0.3	349980	6	ABQ81847	Human DNA
C 860	18	0.3	15254	4	AAK83135	AAK83135 Human imm	C 933	18	0.3	349980	6	ABQ81847	Human DNA
C 861	18	0.3	22873	9	ADCB7198	ADCB7198 Human GPC	C 934	18	0.3	349980	6	ABQ81847	Human DNA
C 862	18	0.3	22873	9	ADCB7198	ADCB7198 Human GPC	C 935	18	0.3	349980	6	ABQ81847	Human DNA
C 863	18	0.3	28315	9	ADCB6254	ADCB6254 Human GPC	C 936	18	0.3	349980	6	ABQ81847	Human DNA
C 864	18	0.3	30610	5	ABAI1643	AbAI1643 Human ner	C 937	18	0.3	349980	6	ABQ81847	Human DNA
C 865	18	0.3	31871	8	ADBS60672	ADBS60672 Connectiv	C 938	18	0.3	349980	6	ABQ81847	Human DNA
C 866	18	0.3	31871	8	ADBS60672	ADBS60672 Connectiv	C 939	18	0.3	349980	6	ABQ81847	Human DNA
C 867	18	0.3	32200	8	ABK42917	ABK42917 Genomic s	C 940	18	0.3	349980	6	ABQ81847	Human DNA
C 868	18	0.3	32200	8	ABK42917	ABK42917 Genomic s	C 941	18	0.3	349980	6	ABQ81847	Human DNA
C 869	18	0.3	34269	4	AAK85168	AAK85168 Human imm	C 942	18	0.3	349980	6	ABQ81847	Human DNA
C 870	18	0.3	34269	4	AAK85168	AAK85168 Human imm	C 943	18	0.3	349980	6	ABQ81847	Human DNA
C 871	18	0.3	36776	4	AAK85321	AAK85321 Human imm	C 944	18	0.3	349980	6	ABQ81847	Human DNA
C 872	18	0.3	36776	4	AAK85321	AAK85321 Human imm	C 945	18	0.3	349980	6	ABQ81847	Human DNA
C 873	18	0.3	36643	9	ADBS72749	ADBS72749 Mouse NCF	C 946	18	0.3	349980	6	ABQ81847	Human DNA
C 874	18	0.3	36643	9	ADBS72749	ADBS72749 Mouse NCF	C 947	18	0.3	349980	6	ABQ81847	Human DNA
C 875	18	0.3	40328	4	AAZ92584	AAZ92584 Human DAZ	C 948	18	0.3	349980	6	ABQ81847	Human DNA
C 876	18	0.3	40328	4	AAZ92584	AAZ92584 Human DAZ	C 949	18	0.3	349980	6	ABQ81847	Human DNA
C 877	18	0.3	40742	4	AAK79886	AAK79886 Human imm	C 950	18	0.3	349980	6	ABQ81847	Human DNA
C 878	18	0.3	40742	4	AAK79886	AAK79886 Human imm	C 951	18	0.3	349980	6	ABQ81847	Human DNA
C 879	18	0.3	43795	3	AAZ92583	AAZ92583 Human DAZ	C 952	18	0.3	349980	6	ABQ81847	Human DNA
C 880	18	0.3	43795	3	AAZ92583	AAZ92583 Human DAZ	C 953	18	0.3	349980	6	ABQ81847	Human DNA
C 881	18	0.3	45000	9	ADBS9873	ADBS9873 Partial g	C 954	18	0.3	349980	6	ABQ81847	Human DNA
C 882	18	0.3	47670	4	ABLI16824	AbLI16824 Drosophila	C 955	18	0.3	349980	6	ABQ81847	Human DNA
C 883	18	0.3	50368	4	ABLI16768	AbLI16768 Drosophila	C 956	18	0.3	349980	6	ABQ81847	Human DNA
C 884	18	0.3	53522	6	AAZ30228	AAZ30228 Human PKD	C 957	18	0.3	349980	6	ABQ81847	Human DNA
C 885	18	0.3	53522	6	AAZ30228	AAZ30228 Human PKD	C 958	18	0.3	349980	6	ABQ81847	Human DNA
C 886	18	0.3	53577	2	AAAT1851	AAAT1851 Human PKD	C 959	18	0.3	349980	6	ABQ81847	Human DNA
C 887	18	0.3	53577	2	AAAT1851	AAAT1851 Human PKD	C 960	18	0.3	349980	6	ABQ81847	Human DNA
C 888	18	0.3	67571	7	AAAD53224_5	AAAD53224_5 of	C 961	18	0.3	349980	6	ABQ81847	Human DNA
C 889	18	0.3	67571	7	AAAD53224_5	AAAD53224_5 of	C 962	18	0.3	349980	6	ABQ81847	Human DNA
C 890	18	0.3	67571	7	AAAD53224_5	AAAD53224_5 of	C 963	18	0.3	349980	6	ABQ81847	Human DNA
C 891	18	0.3	67571	7	AAAD53224_5	AAAD53224_5 of	C 964	18	0.3	349980	6	ABQ81847	Human DNA
C 892	18	0.3	67571	7	AAAD53224_5	AAAD53224_5 of	C 965	18	0.3	349980	6	ABQ81847	Human DNA
C 893	18	0.3	67571	7	AAAD53224_5	AAAD53224_5 of	C 966	18	0.3	349980	6	ABQ81847	Human DNA
C 894	18	0.3	67571	7	AAAD53224_5	AAAD53224_5 of	C 967	18	0.3	349980	6	ABQ81847	Human DNA
C 895	18	0.3	67571	7	AAAD53224_5	AAAD53224_5 of	C 968	18	0.3	349980	6	ABQ81847	Human DNA
C 896	18	0.3	67571	7	AAAD53224_5	AAAD53224_5 of	C 969	18	0.3	349980	6	ABQ81847	Human DNA
C 897	18	0.3	67571	7	AAAD53224_5	AAAD53224_5 of	C 970	18	0.3	349980	6	ABQ81847	Human DNA
C 898	18	0.3	67571	7	AAAD53224_5	AAAD53224_5 of	C 971	18	0.3	349980	6	ABQ81847	Human DNA
C 899	18	0.3	67571	7	AAAD53224_5	AAAD53224_5 of	C 972	18	0.3	349980	6	ABQ81847	Human DNA

ALIGNMENTS

PR	06-OCT-2000;	2000US-0238396P.
PR	06-OCT-2000;	2000US-0238399P.
PR	16-MAR-2001;	2001US-0276657P.
PR	31-MAY-2001;	2001US-0294823P.
PR	12-JUL-2001;	2001US-0304866P.
PR	26-SEP-2001;	2001US-00235631.
XX		
XX	(CURA-) CUREGEN CORP.	
PA		
XX	Gerlach VL, Macdonagall JR, Smithson G, Millet I, Stone D,	
PI	Ganther E, Ellemann K, Grosse WM, Alsobrook JP, Lepley DM,	
PI	Burgess CE, Padigar M, Kekuda R, Spytek KA, Leach MD, Shinkens RA;	
PT		
PT	atherosclerosis, diabetes, cancer, Alzheimer's disease, hemophilia and	
XX	stroke.	
PS		
PS	WPI, 2002-499860/53.	
XX	P-FSDB; AB678364.	
XX		
CC	Claim 8, Page 58-59; 308pp; English.	
CC		
CC	The invention discloses the isolated human polypeptides, and	
CC	polynucleotides encoding them, that have been designated NOVX. The	
CC	polynucleotides, polynucleotides and antibodies are useful in treating or	
CC	preventing a NOVX-associated disorder which is cardiomyopathy,	
CC	atherosclerosis and diabetes in a human, where the disorder is related to	
CC	cell signal processing and metabolic pathway modulation. They can also be	
CC	used in determining the presence of, or predisposition to, a disease	
CC	associated with altered levels of the polypeptides and polynucleotides of	
CC	any one of the 13 sequences (NOV1-NOV8), for raising antibodies, for	
CC	identifying an agent that binds to, or that modulates the expression or	
CC	activity of the polypeptide, for identifying an agent which is cellular	
CC	receptor or downstream effector, for treating or preventing a NOVX-	
CC	associated disorder and as a pharmaceutical composition comprising the	
CC	polynucleotides, polynucleotide or the antibody. The polypeptides and	
CC	polynucleotides are useful in diagnostic applications (e.g. as a marker	
CC	for cancerous cells or tissue types) where their amounts are assessed, or	
CC	for the manufacture of a medicament (e.g. gene therapy) for treating or	
CC	preventing disorders or syndromes such as hypertension, congenital heart	
CC	defects, aortic stenosis, obesity, infectious disease, anorexia, cancer,	
CC	Alzheimer's disease, Parkinson's disorders, neurodegenerative disorders,	
CC	hemophilia, dyslipidemias, hematopoietic diseases, scleroderma,	
CC	fertility, idiopathic thrombocytopenic purpura, graft versus host	
CC	disease, Crohn's disease, multiple sclerosis, cirrhosis, autoimmune	
CC	disease, systemic lupus erythematosus, asthma, arthritis, psoriasis,	
CC	allergy, stroke, anxiety, Leash-Nyhan syndrome, schizophrenia, cerebellar	
CC	ataxia, pain and alcoholism. They may also be used as immunogens to	
CC	produce antibodies specific for the invention, and as vaccines.	
CC	Transgenic cells containing a NOVX expressing construct are useful to	
CC	produce non-human transgenic animals for studying the function and/or	
CC	activity of the NOVX proteins and for identifying and/or evaluating	
CC	modulators of NOVX protein activity. Transgenic cells containing a NOVX	
CC	expressing construct are useful to produce non-human transgenic animals	
CC	for studying the function and/or activity of the NOVX proteins and for	
CC	identifying and/or evaluating modulators of NOVX protein activity. The	
CC	sequences presented in AB563431-AB563444 are the human NOV1-NOV8 cDNAs	
XX		
SQ	Sequence 5691 BP; 1318 A; 1658 C; 1574 G; 1141 T; 0 U; 0 Other;	
	Query Match 100.0%; Score 5691; DB 6; Length 5691;	
	Best Local Similarity 100.0%; Pred. No. 0;	
	Matches 5691; Conservative 0; Mismatches 0; Indels 0; Gaps 0	
OY	1 ATGAAGCCATGCGCTTGAAGCTGACCTGCTTTCTCCCACTCCTCATGAGGGGCAAG	60
DB	1 ATGAAGCCATGCGCTTGAAGCTGACCTGCTTTCTCCCACTCCTCATGAGGGGCAAG	60
OY	61 GGGTCTCCACCTTGTCTCACTCCGAGCAGACCCCGGTGTGCCAAGACGCGGTCAATT	120
DB	61 GGGTCTCCACCTTGTCTCACTCCGAGCAGACCCCGGTGTGCCAAGACGCGGTCAATT	120

QY 121 GTCAATTCGAGAGAGAGCCCGGAGAGTTTCATCATCCTGGTGGATGAGAGACA 180
DB 121 GTCAATTCGAGAGAGAGCCCGGAGAGTTTCATCATCCTGGTGGATGAGAGACA 180
QY 181 GGAACATTTACTTGGGGGCGGTCAATCGGATTTACAGACTCTCCAGGACCTGAAGTTC 240
DB 181 GGAACATTTACTTGGGGGCGGTCAATCGGATTTACAGACTCTCCAGGACCTGAAGTTC 240
QY 241 TTGGTGCGCATGAGACAGGGCCGAGAGAGACACCCCAAGTTTAAACCAACCCGAGATC 300
DB 241 TTGGTGCGCATGAGACAGGGCCGAGAGAGACACCCCAAGTTTAAACCAACCCGAGATC 300
QY 301 GTCCAGACTGCAATGAGACCTGACCCACACCAACATATGTCACAAAGATGCTCTCATAT 360
DB 301 GTCCAGACTGCAATGAGACCTGACCCACACCAACATATGTCACAAAGATGCTCTCATAT 360
QY 361 GACTACAGAGAGAAACAGGCTGATTTGCTGGAGGCTGTACCAAGGACCTGCAAGTTCG 420
DB 361 GACTACAGAGAGAAACAGGCTGATTTGCTGGAGGCTGTACCAAGGACCTGCAAGTTCG 420
QY 421 CTGAGGCTGAGAGACCTTCAAGCTGGGGAGGCTTATATAGAGAGGACCTATCTG 480
DB 421 CTGAGGCTGAGAGACCTTCAAGCTGGGGAGGCTTATATAGAGAGGACCTATCTG 480
QY 481 TCAGGTGTCAAAGAGCGGCTCAAGTTTGAAGTATGCTTCTTCAAGCAACTGGAT 540
DB 481 TCAGGTGTCAAAGAGCGGCTCAAGTTTGAAGTATGCTTCTTCAAGCAACTGGAT 540
QY 541 GACAAGCTTCAATTCGACAGGAGTGGAGTGGAGAGCCGAGATTTTCCACCAATCTCC 600
DB 541 GACAAGCTTCAATTCGACAGGAGTGGAGTGGAGAGCCGAGATTTTCCACCAATCTCC 600
QY 601 AGCCGAGAACTGACCAAGAACTCTGAGGCGGATGGCATGTTCCGATACGTTCCATGAT 660
DB 601 AGCCGAGAACTGACCAAGAACTCTGAGGCGGATGGCATGTTCCGATACGTTCCATGAT 660
QY 661 GAGTTCGTGGCTGATGATTTAAGATCCCTTCGACACCTTCAACCAATCCTTGACTT 720
DB 661 GAGTTCGTGGCTGATGATTTAAGATCCCTTCGACACCTTCAACCAATCCTTGACTT 720
QY 721 GATATCTACTATGCTATGTTTGAAGTGGCACTTGTCTATCTTTTGAACCTTCAA 780
DB 721 GATATCTACTATGCTATGTTTGAAGTGGCACTTGTCTATCTTTTGAACCTTCAA 780
QY 781 CTTGAGATGATGTCTCCACAGGCTCCACCAACAGAGAGAGGTGTTATATCCAAAGTTC 840
DB 781 CTTGAGATGATGTCTCCACAGGCTCCACCAACAGAGAGAGGTGTTATATCCAAAGTTC 840
QY 841 GTGAGGCTTTGCAAGAGAGACAGACTTCAACTCTATGTAAGAGGAGCCCATTTGGCTGT 900
DB 841 GTGAGGCTTTGCAAGAGAGACAGACTTCAACTCTATGTAAGAGGAGCCCATTTGGCTGT 900
QY 901 GAGCGCACTGGGATGAGATACCGCTGCTGAGGCTGCTTACTGTCCAAAGCGGGGCTC 960
DB 901 GAGCGCACTGGGATGAGATACCGCTGCTGAGGCTGCTTACTGTCCAAAGCGGGGCTC 960
QY 961 GTGCTTGGCAGAGACCTTGGAGTCCATCCAGATATAGACCTGCTTTCACCGCTTCTCC 1020
DB 961 GTGCTTGGCAGAGACCTTGGAGTCCATCCAGATATAGACCTGCTTTCACCGCTTCTCC 1020
QY 1021 AAGGGCCAGAGACCGGAAATCCCTGATGATCGGCGCTGTGCATCTTCATCTTG 1080
DB 1021 AAGGGCCAGAGACCGGAAATCCCTGATGATCGGCGCTGTGCATCTTCATCTTG 1080
QY 1081 AAGCAGATTAATGACCGGATTAAGAGAGGCTGAGATCTTTTACCGGGGCGAGGAGAG 1140
DB 1081 AAGCAGATTAATGACCGGATTAAGAGAGGCTGAGATCTTTTACCGGGGCGAGGAGAG 1140
QY 1141 CTGAGACTGAGCTGAGCTCAAGGTGAGAGACATCCCTGACAGCACTGCTTAAACCAT 1200
DB 1141 CTGAGACTGAGCTGAGCTCAAGGTGAGAGACATCCCTGACAGCACTGCTTAAACCAT 1200
QY 1201 GACGATTAACCTTGTGGCTGAGATGAATGCTCCCTGGAGAGTGTCCGACATGTGTGCT 1260

DB 1201 GACGATTAACCTTGTGGCTGAGATGAATGCTCCCTGGGAGTGTCCGACATGTGTGCT 1260
QY 1261 GGAATTCCTGCTTACAGAGAGACAGGACCGGATGAGAGTGTGATCATCGATATGTCTAC 1320
DB 1261 GGAATTCCTGCTTACAGAGAGACAGGACCGGATGAGAGTGTGATCATCGATATGTCTAC 1320
QY 1321 AAGAACCACTCTGAGCTTGTGGGACCAAAAGTGGCAAGCTTGAAGAGATCGGGTGTG 1380
DB 1321 AAGAACCACTCTGAGCTTGTGGGACCAAAAGTGGCAAGCTTGAAGAGATCGGGTGTG 1380
QY 1381 GATGAGCCCAAGGGGACAGCCCTTCAGTATGAGACGTTGACAGGTGTGGACCCCGGCCCA 1440
DB 1381 GATGAGCCCAAGGGGACAGCCCTTCAGTATGAGACGTTGACAGGTGTGGACCCCGGCCCA 1440
QY 1441 GTCCCTCCGGGATATGAGCTTCTCCAAGGACACAGAGCACTCTACATCATGTCAAGAGAG 1500
DB 1441 GTCCCTCCGGGATATGAGCTTCTCCAAGGACACAGAGCACTCTACATCATGTCAAGAGAG 1500
QY 1501 CAGCTCACAGAGTCCCTGTGAGATCCTGTGTGATGATGAGACTGCGGAGTGCCTT 1560
DB 1501 CAGCTCACAGAGTCCCTGTGAGATCCTGTGTGATGATGAGACTGCGGAGTGCCTT 1560
QY 1561 GGCTCAGGCGACCCCACTGTGGGCTGTGTGCTGCAACAAGTGCACCCGGAAGAGAG 1620
DB 1561 GGCTCAGGCGACCCCACTGTGGGCTGTGTGCTGCAACAAGTGCACCCGGAAGAGAG 1620
QY 1621 CGGTGTGAGCGGATCCAGAGAGCCCGGAGGTTTCTTGGAGATGAACAGATGTGTCCG 1680
DB 1621 CGGTGTGAGCGGATCCAGAGAGCCCGGAGGTTTCTTGGAGATGAACAGATGTGTCCG 1680
QY 1681 CTGAGCGTCAATCCCAATAATATCTCCGCTCTGAGTAAACAGGCTGTGCTGAGAG 1740
DB 1681 CTGAGCGTCAATCCCAATAATATCTCCGCTCTGAGTAAACAGGCTGTGCTGAGAG 1740
QY 1741 ACGTCAATGTCCCGAGACTGTGAGCTGTGAGCTGTCAACTGTCACTTTGAGAGACTGTGAGAG 1800
DB 1741 ACGTCAATGTCCCGAGACTGTGAGCTGTGAGCTGTGAGCTGTCAACTGTCACTTTGAGAGACTGTGAGAG 1800
QY 1801 ATGATGAGGCTGTGTGTGGGCAATGATTCAGATGCTACTCCCTGACAGCCAGAGAGTGTG 1860
DB 1801 ATGATGAGGCTGTGTGTGGGCAATGATTCAGATGCTACTCCCTGACAGCCAGAGAGTGTG 1860
QY 1861 CCCCGGATCATCAAGAGATGGGAGCACCATGTGTGATGAGCTTCAAGTCAAAATCAAAG 1920
DB 1861 CCCCGGATCATCAAGAGATGGGAGCACCATGTGTGATGAGCTTCAAGTCAAAATCAAAG 1920
QY 1921 GAGACCGGCAATGACTTGTGCGAGACACAGCTTGTCTTCAATTTGAGAGCTGTCAAGAT 1980
DB 1921 GAGACCGGCAATGACTTGTGCGAGACACAGCTTGTCTTCAATTTGAGAGCTGTCAAGAT 1980
QY 1981 TCGTGCCTGTCCCTGTGGAGAGTCCATACCGCTGCACTGATGTAATACCGGATGTC 2040
DB 1981 TCGTGCCTGTCCCTGTGGAGAGTCCATACCGCTGCACTGATGTAATACCGGATGTC 2040
QY 2041 TGCACCAATGACCCCAAGACCTGCTCTTCAAGAGGCGGAGTGAAGCTGCGGAGAGAC 2100
DB 2041 TGCACCAATGACCCCAAGACCTGCTCTTCAAGAGGCGGAGTGAAGCTGCGGAGAGAC 2100
QY 2101 TGGCCCCAGCTGCTGAGTGAAGAGATCCGAGGCGGAGAGTGAAGCTGCGGAGAGATC 2160
DB 2101 TGGCCCCAGCTGCTGAGTGAAGAGATCCGAGGCGGAGAGTGAAGCTGCGGAGAGATC 2160
QY 2161 ACGTGAAGGCGCAAGAACTCCCGCAGCCCAAGTGTGGGCAAGCTGTGCTGAATGCAATC 2220
DB 2161 ACGTGAAGGCGCAAGAACTCCCGCAGCCCAAGTGTGGGCAAGCTGTGCTGAATGCAATC 2220
QY 2221 CTCAACATTCAGGAGAGAGAGAGAGAGAGTGCCTGCGCTTCAACAGCTCCAGAGTGA 2280
DB 2221 CTCAACATTCAGGAGAGAGAGAGAGAGAGTGCCTGCGCTTCAACAGCTCCAGAGTGA 2280
QY 2281 CAGTGCAGAACACCTCTTATTTCTATGAAGGATGAGATCAACACCTGCGCTGTGAG 2340

Db 2281 CAGTCCAGAAACCTCTTATTCATGAGGATGAGATCAACAACCTGCCGTGAG 2340
Qy 2341 TTGACAGTGTGTGAAATGGGCACTTCAACATTGACAAACCGCTCAGATTAAGTTAC 2400
Db 2341 TTGACAGTGTGTGAAATGGGCACTTCAACATTGACAAACCGCTCAGATTAAGTTAC 2400
Qy 2401 CTCTCAAGTGTGAGGCAATGGGTGAGCTGCGGCTGTGCTCAAGCTGACCCAGAC 2460
Db 2401 CTCTCAAGTGTGAGGCAATGGGTGAGCTGCGGCTGTGCTCAAGCTGACCCAGAC 2460
Qy 2461 TTGCAATGTGTGAGGCAATGGGTGAGCTGCGGCTGTGCTCAAGCTGACCCAGAC 2520
Db 2461 TTGCAATGTGTGAGGCAATGGGTGAGCTGCGGCTGTGCTCAAGCTGACCCAGAC 2520
Qy 2521 CAGGAGAGCGAGTGTGAGGCAATGGGTGAGCTGCGGCTGTGCTCAAGCTGACCCAGAC 2580
Db 2521 CAGGAGAGCGAGTGTGAGGCAATGGGTGAGCTGCGGCTGTGCTCAAGCTGACCCAGAC 2580
Qy 2581 ACAGAGATATCCCGGTGACAGGCCCCCGGAAAGGGGGCAACCAAGTCACTATCCGAGG 2640
Db 2581 ACAGAGATATCCCGGTGACAGGCCCCCGGAAAGGGGGCAACCAAGTCACTATCCGAGG 2640
Qy 2641 GAGAACCTGGGCTGTGAAATTTGCGCAATGCTCCATGTCAGAGTGTGAGATGGGGAG 2700
Db 2641 GAGAACCTGGGCTGTGAAATTTGCGCAATGCTCCATGTCAGAGTGTGAGATGGGGAG 2700
Qy 2701 TGCAGCCCTTTAGTGTGATGATTCATCCCTGCAAGACAGATGCTGTGAGATGGGGAG 2760
Db 2701 TGCAGCCCTTTAGTGTGATGATTCATCCCTGCAAGACAGATGCTGTGAGATGGGGAG 2760
Qy 2761 GCCAAGCCCAAGCAGATGAGGCTTGTGAGATGCTGCTGCTGTGTGTGCTGAA 2820
Db 2761 GCCAAGCCCAAGCAGATGAGGCTTGTGAGATGCTGCTGCTGTGTGTGCTGAA 2820
Qy 2821 TTGATGGCCCGGCTGTGAGATGATTCATCCCTGCAAGACAGATGCTGTGAGATGGGGAG 2880
Db 2821 TTGATGGCCCGGCTGTGAGATGATTCATCCCTGCAAGACAGATGCTGTGAGATGGGGAG 2880
Qy 2881 CCCAGCCGGGGCCCATGTGCGGAGGAGCCCAAGTCACTAATGAAAGTGTG 2940
Db 2881 CCCAGCCGGGGCCCATGTGCGGAGGAGCCCAAGTCACTAATGAAAGTGTG 2940
Qy 2941 GCGGAGAGCAAGTGTGAGATGATTCATCCCTGCAAGACAGATGCTGTGAGATGGGGAG 3000
Db 2941 GCGGAGAGCAAGTGTGAGATGATTCATCCCTGCAAGACAGATGCTGTGAGATGGGGAG 3000
Qy 3001 CCAATCTCAATTTGCTGCAACACCAATCCTCAGATGAGTGTGAGATGAAAGTGTG 3060
Db 3001 CCAATCTCAATTTGCTGCAACACCAATCCTCAGATGAGTGTGAGATGAAAGTGTG 3060
Qy 3061 GTGAGGTGTGAGGAGGCAAGATTCACAGAGACCTGCTTTGAGATGTGAAAGCCCC 3120
Db 3061 GTGAGGTGTGAGGAGGCAAGATTCACAGAGACCTGCTTTGAGATGTGAAAGCCCC 3120
Qy 3121 ACCATCGTGTGAGATGAGGCAAGATGAGATGCTGAGTGTGAAAGCAACCCATCGCGTA 3180
Db 3121 ACCATCGTGTGAGATGAGGCAAGATGAGATGCTGAGTGTGAAAGCAACCCATCGCGTA 3180
Qy 3181 TGGGGAGCCCACTGTGACCTTCATACAGAACCCCAATCGGTGTGCAAGCATGAGGAGAG 3240
Db 3181 TGGGGAGCCCACTGTGACCTTCATACAGAACCCCAATCGGTGTGCAAGCATGAGGAGAG 3240
Qy 3241 GAGCAATCAATATCTGTGAGTGTGAGATGAGATGCTGAGTGTGAGGCGCGCGC 3300
Db 3241 GAGCAATCAATATCTGTGAGTGTGAGATGAGATGCTGAGTGTGAGGCGCGCGC 3300
Qy 3301 CTGCTCTGGGTCTGTACCAACAGTCAAGCTGACCGAGAGGCGCGAGAGATTGGCTTC 3360
Db 3301 CTGCTCTGGGTCTGTACCAACAGTCAAGCTGACCGAGAGGCGCGAGAGATTGGCTTC 3360
Qy 3361 ATCTGTGACAAAGTGTGAGTGTGAGATGAGATGCTGAGTGTGAGGCGCGCGC 3420
Db 3361 ATCTGTGACAAAGTGTGAGTGTGAGATGAGATGCTGAGTGTGAGGCGCGCGC 3420

Qy 3421 AACCGGTGTTGAGGCTTTGGTCCCTCAGAGATCTGAGAGCTCAAGCTGAGCGCC 3480
Db 3421 AACCGGTGTTGAGGCTTTGGTCCCTCAGAGATCTGAGAGCTCAAGCTGAGCGCC 3480
Qy 3481 ATCATCTTAAAGGCAAGAACTGATCCGCTGTGAGTGTGAGGAGGCAAGTGAAGTTAAC 3540
Db 3481 ATCATCTTAAAGGCAAGAACTGATCCGCTGTGAGTGTGAGGAGGCAAGTGAAGTTAAC 3540
Qy 3541 TACAGTGTGTGTTGGGAGAGAACCGGTGACCGGTGACCGGTGACAGATGTGAGCTGTC 3600
Db 3541 TACAGTGTGTGTTGGGAGAGAACCGGTGACCGGTGACCGGTGACAGATGTGAGCTGTC 3600
Qy 3601 TGGAGTGTGAGGCAAGTGTGAGGCAAGTGTGAGGCAAGTGTGAGGCAAGTGTGAGG 3660
Db 3601 TGGAGTGTGAGGCAAGTGTGAGGCAAGTGTGAGGCAAGTGTGAGGCAAGTGTGAGG 3660
Qy 3661 TACTCCCGGGGATGCTGTACATTTGCCCCGAGACAGCCGCTACAGCTGCGCGCATGTC 3720
Db 3661 TACTCCCGGGGATGCTGTACATTTGCCCCGAGACAGCCGCTACAGCTGCGCGCATGTC 3720
Qy 3721 AGCATGCAATGTGCTGAGGCTGCTCTCATATTTTCATGCTGAGCCGTGCTCATTTGCTAT 3780
Db 3721 AGCATGCAATGTGCTGAGGCTGCTCTCATATTTTCATGCTGAGCCGTGCTCATTTGCTAT 3780
Qy 3781 AAACGCAAGTCCCGCAAGTGTGAGGCAAGTGTGAGGCAAGTGTGAGGCAAGTGTGAGG 3840
Db 3781 AAACGCAAGTCCCGCAAGTGTGAGGCAAGTGTGAGGCAAGTGTGAGGCAAGTGTGAGG 3840
Qy 3841 CTGAGTGTGAGGCTGAGTGTGAGGCAAGTGTGAGGCAAGTGTGAGGCAAGTGTGAGG 3900
Db 3841 CTGAGTGTGAGGCTGAGTGTGAGGCAAGTGTGAGGCAAGTGTGAGGCAAGTGTGAGG 3900
Qy 3901 CATGAGTGTGAGGCTGAGTGTGAGGCAAGTGTGAGGCAAGTGTGAGGCAAGTGTGAGG 3960
Db 3901 CATGAGTGTGAGGCTGAGTGTGAGGCAAGTGTGAGGCAAGTGTGAGGCAAGTGTGAGG 3960
Qy 3961 AACATGAGGAGTGTGAGGCAAGTGTGAGGCAAGTGTGAGGCAAGTGTGAGGCAAGTGTGAGG 4020
Db 3961 AACATGAGGAGTGTGAGGCAAGTGTGAGGCAAGTGTGAGGCAAGTGTGAGGCAAGTGTGAGG 4020
Qy 4021 CCGGAGTGTGAGGCAAGTGTGAGGCAAGTGTGAGGCAAGTGTGAGGCAAGTGTGAGGCAAGTGTGAGG 4080
Db 4021 CCGGAGTGTGAGGCAAGTGTGAGGCAAGTGTGAGGCAAGTGTGAGGCAAGTGTGAGGCAAGTGTGAGG 4080
Qy 4081 AACAGGAGTGTGAGGCAAGTGTGAGGCAAGTGTGAGGCAAGTGTGAGGCAAGTGTGAGGCAAGTGTGAGG 4140
Db 4081 AACAGGAGTGTGAGGCAAGTGTGAGGCAAGTGTGAGGCAAGTGTGAGGCAAGTGTGAGGCAAGTGTGAGG 4140
Qy 4141 GCGAGCGTGTGAGGCAAGTGTGAGGCAAGTGTGAGGCAAGTGTGAGGCAAGTGTGAGGCAAGTGTGAGG 4200
Db 4141 GCGAGCGTGTGAGGCAAGTGTGAGGCAAGTGTGAGGCAAGTGTGAGGCAAGTGTGAGGCAAGTGTGAGG 4200
Qy 4201 GCGAGCGTGTGAGGCAAGTGTGAGGCAAGTGTGAGGCAAGTGTGAGGCAAGTGTGAGGCAAGTGTGAGG 4260
Db 4201 GCGAGCGTGTGAGGCAAGTGTGAGGCAAGTGTGAGGCAAGTGTGAGGCAAGTGTGAGGCAAGTGTGAGG 4260
Qy 4261 AACACCGTGTGAGGCAAGTGTGAGGCAAGTGTGAGGCAAGTGTGAGGCAAGTGTGAGGCAAGTGTGAGG 4320
Db 4261 AACACCGTGTGAGGCAAGTGTGAGGCAAGTGTGAGGCAAGTGTGAGGCAAGTGTGAGGCAAGTGTGAGG 4320
Qy 4321 TGGTTTACTTCTCTCTCTCAAAATTTGCTGAGAGTGTGAGGAGGCGCGCTTCTTC 4380
Db 4321 TGGTTTACTTCTCTCTCTCAAAATTTGCTGAGAGTGTGAGGAGGCGCGCTTCTTC 4380
Qy 4381 CTGTTCTGTGCAATCAAGCAAGTGTGAGGCAAGTGTGAGGCAAGTGTGAGGCAAGTGTGAGGCAAGTGTGAGG 4440
Db 4381 CTGTTCTGTGCAATCAAGCAAGTGTGAGGCAAGTGTGAGGCAAGTGTGAGGCAAGTGTGAGGCAAGTGTGAGG 4440
Qy 4441 GCGGAGTGTGAGGCAAGTGTGAGGCAAGTGTGAGGCAAGTGTGAGGCAAGTGTGAGGCAAGTGTGAGGCAAGTGTGAGG 4500
Db 4441 GCGGAGTGTGAGGCAAGTGTGAGGCAAGTGTGAGGCAAGTGTGAGGCAAGTGTGAGGCAAGTGTGAGGCAAGTGTGAGG 4500

4501 GTCTGAGCTGTGTGACCCAGCAATGGCAACAGCCCGAGGCTCCAGTAAAGATCCTC 4560
 Db 4501 GTCTGAGCTGTGTGACCCAGCAATGGCAACAGCCCGAGGCTCCAGTAAAGATCCTC 4560
 Qy 4561 AACTGTGACCACTGCTGAGGTCAAGGAGAGATTTCTGGATGGCATCTTCAAGATG 4620
 Db 4561 AACTGTGACCACTGCTGAGGTCAAGGAGAGATTTCTGGATGGCATCTTCAAGATG 4620
 Qy 4621 CCTTGTCCACCCGCGCCAAAGCTGAGATATGATCTTGGATGGGCAAGAAAGTGG 4680
 Db 4621 CCTTGTCCACCCGCGCCAAAGCTGAGATATGATCTTGGATGGGCAAGAAAGTGG 4680
 Qy 4681 GCAGAGATGATCTTGGAGATGAAAGATCAACCAACCAATTGAAAGATGGAAGCA 4740
 Db 4681 GCAGAGATGATCTTGGAGATGAAAGATCAACCAACCAATTGAAAGATGGAAGCA 4740
 Qy 4741 CTGAGCACTGAGCCACTACAGAGTGCAGATGATTCGGTGGTGGCATTTAGTGTCAAG 4800
 Db 4741 CTGAGCACTGAGCCACTACAGAGTGCAGATGATTCGGTGGTGGCATTTAGTGTCAAG 4800
 Qy 4801 CAGGTGACAGCTTATTAACGAGTGAACAACTCCAGTCTCCAGAGCTCAGCAAGTAA 4860
 Db 4801 CAGGTGACAGCTTATTAACGAGTGAACAACTCCAGTCTCCAGAGCTCAGCAAGTAA 4860
 Qy 4861 TATGAAACATGATCCGGTACAGGGCAGCCGAGCTCCGCTCAGGACACTATG 4920
 Db 4861 TATGAAACATGATCCGGTACAGGGCAGCCGAGCTCCGCTCAGGACACTATG 4920
 Qy 4921 ATCACTCTGACCTGAGAGTGAAGTCAAGATGTGGACCTTAGTGAAGCAACAGCAC 4980
 Db 4921 ATCACTCTGACCTGAGAGTGAAGTCAAGATGTGGACCTTAGTGAAGCAACAGCAC 4980
 Qy 4981 GGAGACCAAGAGAGGGGGGACCGGGGAGAGAGATGATGTAATCTACCTGACCCGA 5040
 Db 4981 GGAGACCAAGAGAGGGGGGACCGGGGAGAGAGATGATGTAATCTACCTGACCCGA 5040
 Qy 5041 CTCTGGCCACTAAGGGCACTGACAGAAATTTGTGATGACCTCTTTGAGCAACTCTTC 5100
 Db 5041 CTCTGGCCACTAAGGGCACTGACAGAAATTTGTGATGACCTCTTTGAGCAACTCTTC 5100
 Qy 5101 AGCAGCGCAACCTGGCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 5160
 Db 5101 AGCAGCGCAACCTGGCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 5160
 Qy 5161 GATGAGCAGGCTGATTAACATGAGCTTCAATGACCCGACGTCGGCATCTGGAAGAGC 5220
 Db 5161 GATGAGCAGGCTGATTAACATGAGCTTCAATGACCCGACGTCGGCATCTGGAAGAGC 5220
 Qy 5221 AATTGCTGCTGCTGAGGTTTGGGTCAATGATCAAGAACCCGAGATTTGTGTTGAC 5280
 Db 5221 AATTGCTGCTGCTGAGGTTTGGGTCAATGATCAAGAACCCGAGATTTGTGTTGAC 5280
 Qy 5281 ATTCATTAAGAACAGCATCAAGAGCGCTGCTCTGCTGCTGCTGCTGCTGCTGCTG 5340
 Db 5281 ATTCATTAAGAACAGCATCAAGAGCGCTGCTCTGCTGCTGCTGCTGCTGCTGCTG 5340
 Qy 5341 TCTTGTCCACGTCAGAGCAACCGGTGGGCAAGGATCTGCTGCTGCTGCTGCTGCTG 5400
 Db 5341 TCTTGTCCACGTCAGAGCAACCGGTGGGCAAGGATCTGCTGCTGCTGCTGCTGCTG 5400
 Qy 5401 GCCAAGGACATCCCGAGCTACAGAAATTTGGGTGAGAGATTTACTCAGACATAGGGAAG 5460
 Db 5401 GCCAAGGACATCCCGAGCTACAGAAATTTGGGTGAGAGATTTACTCAGACATAGGGAAG 5460
 Qy 5461 ATGCGAGCCATCAGCGACCAAGACATGAAACGATACCTGCTGAGAGATCCCGAGATG 5520
 Db 5461 ATGCGAGCCATCAGCGACCAAGACATGAAACGATACCTGCTGAGAGATCCCGAGATG 5520
 Qy 5521 ATGAAATGATTAACAACATGAGTGAAGTCTCAGAGATCTTCTCCTAATGAGGCAATAG 5580
 Db 5521 ATGAAATGATTAACAACATGAGTGAAGTCTCAGAGATCTTCTCCTAATGAGGCAATAG 5580
 Qy 5581 AGCGAGAGATCTTGGACCTCTGAGACCAAGATGACCAAGTGTGGAGAACCAATTGGCC 5640

Db 5581 AGCGAGAGATCTTGGACCTCTGAGACCAAGATGACCAAGTGTGGAGAACCAATTGGCC 5640
 Qy 5641 TACAACTAGAACAAAGTCATTAACCTCATAGCTTAGACAGCTGAATATA 5691
 Db 5641 TACAACTAGAACAAAGTCATTAACCTCATAGCTTAGACAGCTGAATATA 5691
 RESULT 2
 ABA00062
 ID ABA00062 standard; cDNA; 6367 BP.
 XX ABA00062;
 AC
 XX 25-OCT-2002 (first entry)
 DE CADHP-9 coding sequence, Incyte ID No: 7156379CBI.
 XX
 DE Gene; human; cell adhesion protein; CADHP; AIDS; Alzheimer's disease;
 KW acquired immunodeficiency syndrome; thymic dysplasia; epilepsy;
 KW renal tubular acidosis; congenital glaucoma; cancer; atherosclerosis;
 KW Parkinson's disease; ss.
 XX Homo sapiens.
 OS
 FH Key Location/Qualifiers
 FT 566..6250
 CDS /*tag= a
 FT /product= "CADHP-9"
 FT
 FN WO200259312-A2.
 XX
 PD 01-AUG-2002.
 XX
 PF 18-DEC-2001; 2001WO-US049206.
 XX
 PR 18-DEC-2000; 2000US-0256542P.
 PR 22-DEC-2000; 2000US-0259604P.
 PR 05-JAN-2001; 2001US-0260101P.
 XX
 PA (INCYTE GENOMICS INC.
 XX
 PI Duggan BM, Xu Y, Lee EA, Lee S, Lu DM, Warren BA, Yue H;
 PI Gierzen KJ, Horchell CD, Burford N, Baughn MR, Tang TY, Hillman JL;
 PI Gandhi AR, Kallik DA, Bandman O, Graul RC, Walla NK, Lu Y;
 PI Ramkumar J, Yao MG, Lai PG;
 XX
 DR WPI; 2002-590826/63.
 DR P-PSDB; AAG79420.
 XX
 PT New human cell adhesion proteins (CADHP) useful for treating, diagnosing
 PT and preventing diseases or conditions associated with the aberrant CADHP
 PT expression e.g. cancer, acquired immunodeficiency syndrome, Alzheimer's
 PT disease and epilepsy.
 XX
 PS Claim 5; Page 146-48; 149pp; English.
 XX
 CC The sequences given in ABA00054-63 encode novel human cell adhesion
 CC proteins (CADHP). The CADHP polypeptides and polynucleotides are useful
 CC in treating, diagnosing and preventing diseases or conditions associated
 CC with the decreased expression or overexpression of CADHP, e.g. immune
 CC system (acquired immunodeficiency syndrome, thymic dysplasia),
 CC neurological (Alzheimer's disease, Parkinson's disease, epilepsy),
 CC developmental (renal tubular acidosis, congenital glaucoma) and cell
 CC proliferative (cancer, atherosclerosis) disorders. They are also useful
 CC in assessing the effects of exogenous compounds on the expression of
 CC nucleic acid and amino acid sequences of CADHP. The CADHP or its
 CC fragments are useful in screening compounds for effectiveness as agonist
 CC or antagonist of the polypeptides, or in altering the expression of the
 CC target polynucleotide and compounds that specifically bind to or modulate
 CC the activity of the polypeptide. The protein encoded by this cDNA
 CC sequence shows homology to mouse plexin-2
 XX

Sequence 6367 BP; 1423 A; 1885 C; 1799 G; 1259 T; 0 U; 1 Other;
Query Match 97.2%; Score 5532; DB 6; Length 6367;
Best Local Similarity 99.9%; Pred. No. 0;
Matches 5682; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 AAAAAAGCATGCGCTTGGAACTGGACCTGCTTCTCTCCACCTCTGATGCTGGGACATG 60
DB 566 AAAAAAGCATGCGCTTGGAACTGGACCTGCTTCTCTCCACCTCTGATGCTGGGACATG 625
QY 61 GCGTCTCCACTTGTGCTCAACCGGAGCCGAGCCGCTGTCCAGAGAGAGCGGTCAATT 120
DB 626 GCGTCTCCACTTGTGCTCAACCGGAGCCGAGCCGCTGTCCAGAGAGAGCGGTCAATT 685
QY 121 GTCACTTCCGAGAGAGAGCCCGCCGAGGCTTCAATCACTGTGTGTGAGAGAGCA 180
DB 686 GTCACTTCCGAGAGAGAGCCCGCCGAGGCTTCAATCACTGTGTGTGAGAGAGCA 745
QY 181 GGAACATTTACTTGGGGCCGCTCAATCGATTTTCAAGCTCTCCAGAGAGCTGAGGTC 240
DB 746 GGAACATTTACTTGGGGCCGCTCAATCGATTTTCAAGCTCTCCAGAGAGCTGAGGTC 805
QY 241 TTGGTGAAGCATGAGACAGGGCCGAGAGACAAACCCCAAGTGTATCCACCCCGCATC 300
DB 806 TTGGTGAAGCATGAGACAGGGCCGAGAGACAAACCCCAAGTGTATCCACCCCGCATC 865
QY 301 GTCCAGACCTCGCAATGAGCCCTGACCAACCAACATGTCAACAAAGATGCTCTCATTA 360
DB 866 GTCCAGACCTCGCAATGAGCCCTGACCAACCAACATGTCAACAAAGATGCTCTCATTA 925
QY 361 GACTCAAGAGAGAAAGGCTGATGCTGTGGAGCTTATCCAAAGGATCTGCAAGTCG 420
DB 926 GACTCAAGAGAGAAAGGCTGATGCTGTGGAGCTTATCCAAAGGATCTGCAAGTCG 985
QY 421 CTGAGGCTGAGAGAGCTCTTCAAGCTGGGGAGACCTTATCAAGAAGAGACATATCTG 480
DB 986 CTGAGGCTGAGAGAGCTCTTCAAGCTGGGGAGACCTTATCAAGAAGAGACATATCTG 1045
QY 481 TCAGGTCTCAACAGAGACGGCTCAATCTTGAAGTATGCTCTCAACGAACCTGAT 540
DB 1046 TCAGGTCTCAACAGAGACGGCTCAATCTTGAAGTATGCTCTCAACGAACCTGAT 1105
QY 541 GACAAGCTGTCTATTCGACCGGAGTGAATGGGAACCGAGATTTTCCACCAATCTCC 600
DB 1106 GACAAGCTGTCTATTCGACCGGAGTGAATGGGAACCGAGATTTTCCACCAATCTCC 1165
QY 601 AGCCGGAACTGACCAAGAACTCTGAGCGAGTGGATGTCGCTAGCTTCTTCCATGAT 660
DB 1166 AGCCGGAACTGACCAAGAACTCTGAGCGAGTGGATGTCGCTAGCTTCTTCCATGAT 1225
QY 661 GAAGTCTGTGCTGTGATGATTAAGTATCCCTTCCGACACCTTCAATATCCCTGACTT 720
DB 1226 GAAGTCTGTGCTGTGATGATTAAGTATCCCTTCCGACACCTTCAATATCCCTGACTT 1285
QY 721 GATATCTACTATGTCTATGTTTGAAGTGGCAACTTTGTCTAATTTTGAACCTTCGA 780
DB 1286 GATATCTACTATGTCTATGTTTGAAGTGGCAACTTTGTCTAATTTTGAACCTTCGA 1345
QY 781 CCTGAGATGCTGTCTCAACAGGCTCCACCAACCAAGAGAGAGTGTATATCTCAAGTC 840
DB 1346 CCTGAGATGCTGTCTCAACAGGCTCCACCAACCAAGAGAGAGTGTATATCTCAAGTC 1405
QY 841 GTGAGGCTTTGCAAGAGAGACAGCTTCAACTCTATGTAGAGTGGCCATTTGCTGT 900
DB 1406 GTGAGGCTTTGCAAGAGAGACAGCTTCAACTCTATGTAGAGTGGCCATTTGCTGT 1465
QY 901 GAGGCGAGTGGGCTGAGATACCGCTGTGAGGCTGTACTCTGTCCAAAGCGGGGCTC 960
DB 1466 GAGGCGAGTGGGCTGAGATACCGCTGTGAGGCTGTACTCTGTCCAAAGCGGGGCTC 1525
QY 961 GTGCTGGCAGAGACCTTGGAGTGCATCAAGATGATGACTGTCTTCAACGCTTCTCC 1020
DB 1526 GTGCTGGCAGAGACCTTGGAGTGCATCAAGATGATGACTGTCTTCAACGCTTCTCC 1585

QY 1021 AAGGCGCAGAAAGCCGAAATGAAATATCCCTGATGATGCGCCCTGTGCACTTCAATCTTG 1080
DB 1586 AAGGCGCAGAAAGCCGAAATGAAATATCCCTGATGATGCGCCCTGTGCACTTCAATCTTG 1645
QY 1081 AAGCAGATTAATGACCGCATTAAGAGCCGCTGCACTTGTTAACCGGGGCGAGGACG 1140
DB 1646 AAGCAGATTAATGACCGCATTAAGAGCCGCTGCACTTGTTAACCGGGGCGAGGACG 1705
QY 1141 CTGAGCTGTGCTGTGCTCAAGTGAAGAGATATCCCTGAGAGAGTGGCTTAAACAT 1200
DB 1706 CTGAGCTGTGCTGTGCTCAAGTGAAGAGATATCCCTGAGAGAGTGGCTTAAACAT 1765
QY 1201 GACGATTAATCTGTGAGCTGGAATGATGCTCCCTGGAGTGTCCGACATGTGCGT 1260
DB 1766 GACGATTAATCTGTGAGCTGGAATGATGCTCCCTGGAGTGTCCGACATGTGCGT 1825
QY 1261 GGAATTCGCTCTTCAAGAGAGACAGGACCGCATGACGTCTGTCAATGCGATATGCTAC 1320
DB 1826 GGAATTCGCTCTTCAAGAGAGACAGGACCGCATGACGTCTGTCAATGCGATATGCTAC 1885
QY 1321 AAGAACCACTCTGTGAGCTTTGTGGGCAACCAAGTGGCAAGCTGGAAGAAATCCGAGT 1380
DB 1886 AAGAACCACTCTGTGAGCTTTGTGGGCAACCAAGTGGCAAGCTGGAAGAAATCCGAGT 1945
QY 1381 GATGAGACCAAGGAGCAACGCTCTCCAGATGAGACGCTGAGAGTGGTGGACCCCGGCTCA 1440
DB 1946 GATGAGACCAAGGAGCAACGCTCTCCAGATGAGACGCTGAGAGTGGTGGACCCCGGCTCA 2005
QY 1441 GTCCCTGGGAGATATGAGCTTCTCCAAAGAGACCAAGAGAACTCAATCATGTCAAGAG 1500
DB 2006 GTCCCTGGGAGATATGAGCTTCTCCAAAGAGACCAAGAGAACTCAATCATGTCAAGAG 2065
QY 1501 CAGCTCAACAGAGTCTCTGTGAGTCTGTCTGTCTGTCTGTCTGTCTGTCTGTCTGTCT 1560
DB 2066 CAGCTCAACAGAGTCTCTGTGAGTCTGTCTGTCTGTCTGTCTGTCTGTCTGTCTGTCT 2125
QY 1561 GGTCTAAGGACACCCCACTGT 1620
DB 2126 GGTCTAAGGACACCCCACTGT 2185
QY 1621 CGGTGTAGCGGTCAAGAGAGAGCCCGAGGTTTGCCTGGAGATGAAGCAGTGTGTCCG 1680
DB 2245 CGGTGTAGCGGTCAAGAGAGAGCCCGAGGTTTGCCTGGAGATGAAGCAGTGTGTCCG 2245
QY 1681 CTGACGCTCCATCCCAATATCTCGGTCTCAAGTCAAGTGTGTGTGTGTGTGTGTGTGTGT 1740
DB 2246 CTGACGCTCCATCCCAATATCTCGGTCTCAAGTCAAGTGTGTGTGTGTGTGTGTGTGTGT 2305
QY 1741 ACGTACAAATGCTCCGAGAGCTGTCAAGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1800
DB 2306 ACGTACAAATGCTCCGAGAGCTGTCAAGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 2365
QY 1801 ATGAGTGGGCTGT 1860
DB 2366 ATGAGTGGGCTGT 2425
QY 1861 CCCCGATCATCAAGAGAAATGGGAGCAACCATGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1920
DB 2426 CCCCGATCATCAAGAGAAATGGGAGCAACCATGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 2485
QY 1921 GAGACCGGATGACCTTGTGCGAGACCAAGCTTGTCTTCAATATGAGAGTGTGTGTGTGTGT 1980
DB 2486 GAGACCGGATGACCTTGTGCGAGACCAAGCTTGTCTTCAATATGAGAGTGTGTGTGTGTGTGT 2545
QY 1981 TGTGCTGT 2040
DB 2546 TGTGCTGT 2605
QY 2041 TGCACCCATGACCCCAAGACTGTCTCTTCAAGAAAGCCGAGTGAAGTGTGTGTGTGTGTGTGT 2100
DB 2606 TGCACCCATGACCCCAAGACTGTCTCTTCAAGAAAGCCGAGTGAAGTGTGTGTGTGTGTGTGT 2665

OY 2101 TCCCCCAGCTCTGCGAGTGGACAAAGATCTGTGTGCCGTGGAGGTGATCAAGCTTATC 2160
Db 2666 TCCCCCAGCTCTGCGAGTGGACAAAGATCTGTGTGCCGTGGAGGTGATCAAGCTTATC 2725
OY 2161 AGCTTGAAGCCAGAACTCTCCCGAGCCCGAGTCTGGGACGCGGGCTACGAAATGCATC 2220
Db 2726 AGCTTGAAGCCAGAACTCTCCCGAGCCCGAGTCTGGGACGCGGGCTACGAAATGCATC 2785
OY 2221 CTCAACATTCAGGGCAGCGAGCAGCGAGTCCCGCCTTGCCTTCAACAGCTCCAGCGTA 2280
Db 2786 CTCAACATTCAGGGCAGCGAGCAGCGAGTCCCGCCTTGCCTTCAACAGCTCCAGCGTA 2845
OY 2281 CAGTGGCAGAACCTCTTATTTCTTATGAAAGGATGGAGATCAACACTGCGCGTGGAG 2340
Db 2846 CAGTGGCAGAACCTCTTATTTCTTATGAAAGGATGGAGATCAACACTGCGCGTGGAG 2905
OY 2341 TTGACAGTGTGTGGAATGGGCACTTCAACATTCACACCCAGCTCAGAAATAAGTTGAC 2400
Db 2906 TTGACAGTGTGTGGAATGGGCACTTCAACATTCACACCCAGCTCAGAAATAAGTTGAC 2965
OY 2401 CTCTAACATGTGTGAGACCATGCTGTGAGAGCTGGCGGCTGTGCTTCAAGGTGACCCAGAC 2460
Db 2966 CTCTAACATGTGTGAGACCATGCTGTGAGAGCTGGCGGCTGTGCTTCAAGGTGACCCAGAC 3025
OY 2461 TTGCGCATGTGGCTGTGGTGGCAGAGGCCAGAGCCAGTGCACCCTGGCAGCACTGCCCTGCC 2520
Db 3026 TTGCGCATGTGGCTGTGGTGGCAGAGGCCAGAGCCAGTGCACCCTGGCAGCACTGCCCTGCC 3085
OY 2521 CAGAGAGCCAGATGAGTGAAGCTGTGTGTGCCAAAGCAAGTGCACAAACCCCGCATC 2580
Db 3086 CAGAGAGCCAGATGAGTGAAGCTGTGTGTGCCAAAGCAAGTGCACAAACCCCGCATC 3145
OY 2581 ACAGAGATATACCCGGTGAACAGGCCCGCCGGGAGAGGGGAGCCAAAGGTCTACTATCCGAGGG 2640
Db 3146 ACAGAGATATACCCGGTGAACAGGCCCGCCGGGAGAGGGGAGCCAAAGGTCTACTATCCGAGGG 3205
OY 2641 GAGAACCTTGGGCTTGAATTTGCGCAATGCGCTCCCATGTCAAGTTGCTGGCGTGGAG 2700
Db 3206 GAGAACCTTGGGCTTGAATTTGCGCAATGCGCTCCCATGTCAAGTTGCTGGCGTGGAG 3265
OY 2701 TGCAGCCCTTTAATGTGATGGTTAATACCCCTGCAGAACAGATGCTGTGTGAGATGGGGAG 2760
Db 3266 TGCAGCCCTTTAATGTGATGGTTAATACCCCTGCAGAACAGATGCTGTGTGAGATGGGGAG 3325
OY 2761 GCCAAGCCACGACGATGACAGGCTTTCGTGAGATCTGCGTGGCTGTGTGTGCGGCTGAA 2820
Db 3326 GCCAAGCCACGACGATGACAGGCTTTCGTGAGATCTGCGTGGCTGTGTGTGCGGCTGAA 3385
OY 2821 TTCAATGCGCGGTCTCTCAACGCTTATTAATTCATGACACTGACTCTCTAGATCTGAAAG 2880
Db 3386 TTCAATGCGCGGTCTCTCAACGCTTATTAATTCATGACACTGACTCTCTAGATCTGAAAG 3445
OY 2881 CCCAGCCGAGGAGCCATGCTCCGAGGAGCCCAAGTGCATCAAGGCAACCACTGAAAT 2940
Db 3446 CCCAGCCGAGGAGCCATGCTCCGAGGAGCCCAAGTGCATCAAGGCAACCACTGAAAT 3505
OY 2941 GCCGGAAAGCAAGTGTGTGTGATTTTGGAAAGGAGGCTGTCTTCTTCCACAGGCAATCT 3000
Db 3506 GCCGGAAAGCAAGTGTGTGTGATTTTGGAAAGGAGGCTGTCTTCTTCCACAGGCAATCT 3565
OY 3001 CCATCTCAATGTCTGTGAACACACACACTCTCAGATAGGTGTGATGAGATGAAAGTGTG 3060
Db 3566 CCATCTCAATGTCTGTGAACACACACTCTCAGATAGGTGTGATGAGATGAAAGTGTG 3625
OY 3061 GTGAGAGTGAACAGGCCCAAGATCAACAGAGACTGTGCTTTCAGTATGTGGAAAGCCC 3120
Db 3626 GTGAGAGTGAACAGGCCCAAGATCAACAGAGACTGTGCTTTCAGTATGTGGAAAGCCC 3685
OY 3121 ACCATGCGGAGATTGAGCCAGAAATGAGCATGTGACGTGGAACACACCATGTGCCGTA 3180
Db 3686 ACCATGCGGAGATTGAGCCAGAAATGAGCATGTGACGTGGAACACACCATGTGCCGTA 3745
OY 3181 TGGGGAGCCACCTGGAATCTATACAAACCCCAAGATCGTGTCCAAAGCATGAGGGAG 3240

Db 3746 TGGGGAGCCACCTGGAATCTATACAAACCCCAAGATCGTGTCCAAAGCATGAGGGAG 3805
OY 3241 GAGACATCAATATCTGTGAGTTCCTGAACGCTATGAGTGAACCTGTCAAGGCGCCGCG 3300
Db 3806 GAGACATCAATATCTGTGAGTTCCTGAACGCTATGAGTGAACCTGTCAAGGCGCCGCG 3865
OY 3301 CTGCTTGTGTCTTGAACCAACAGTGAACCTGACCCGAGAGGCCGAGAGTGTGGCTTC 3360
Db 3866 CTGCTTGTGTCTTGAACCAACAGTGAACCTGACCCGAGAGGCCGAGAGTGTGGCTTC 3925
OY 3361 ATCTGGAACAAGTCCAGTCCCTGCTCATCTCAACAGAACCACTTCACTATGCC 3420
Db 3926 ATCTGGAACAAGTCCAGTCCCTGCTCATCTCAACAGAACCACTTCACTATGCC 3985
OY 3421 AACCCGTTGTTGAGGCTTGTGATCCCTGAGGAATCCTGAGCTCAAGGCTGCAAGCC 3480
Db 3986 AACCCGTTGTTGAGGCTTGTGATCCCTGAGGAATCCTGAGCTCAAGGCTGCAAGCC 4045
OY 3481 ATCATCTTAAAGGCAAGAACTGATCCGCTGTGTGCTGGGCGCAACGTGAAGCTGAAC 3540
Db 4046 ATCATCTTAAAGGCAAGAACTGATCCGCTGTGTGCTGGGCGCAACGTGAAGCTGAAC 4105
OY 3541 TACACTGTGCTGTGGGAGAGAGCCGTGACCCGTGACCGTGTCAAGATGTGCCAGTGTCTC 3600
Db 4106 TACACTGTGCTGTGGGAGAGAGCCGTGACCCGTGACCGTGTCAAGATGTGCCAGTGTCTC 4165
OY 3601 TGGCAGTCCCCCAACTCATCTGCGAGGCAACAAAGTGAATGACCCGTGTGCTGTGATGAG 3660
Db 4166 TGGCAGTCCCCCAACTCATCTGCGAGGCAACAAAGTGAATGACCCGTGTGCTGTGATGAG 4225
OY 3661 TACTCCCCGGGAGATGTGTATATTGCCCCGAGCAGGCCCTCAAGCTTCCCGGCAATGCTC 3720
Db 4226 TACTCCCCGGGAGATGTGTATATTGCCCCGAGCAGGCCCTCAAGCTTCCCGGCAATGCTC 4285
OY 3721 AGCATGCGAGTGGCTGGGGGCTCCTCATATTTTCACTGTGAGCTGATTCCTAT 3780
Db 4286 AGCATGCGAGTGGCTGGGGGCTCCTCATATTTTCACTGTGAGCTGATTCCTAT 4345
OY 3781 AAAGCGAAGTCCCGCAAAAGTGAACCTCAACGCTGAAGCGGCTGAGATGCAATGGAACAC 3840
Db 4346 AAAGCGAAGTCCCGCAAAAGTGAACCTCAACGCTGAAGCGGCTGAGATGCAATGGAACAC 4405
OY 3841 CTGAGATCCGATGTGGGCTCGAGGTGACAGGAACTTTGCGAGCTGACAGCCGATC 3900
Db 4406 CTGAGATCCGATGTGGGCTCGAGGTGACAGGAACTTTGCGAGCTGACAGCCGATC 4465
OY 3901 CATGAGCTGAACAGTGAACCTTGAATGAGCCGGGAATTCGTTCTCTGATCTATAGAACTTAC 3960
Db 4466 CATGAGCTGAACAGTGAACCTTGAATGAGCCGGGAATTCGTTCTCTGATCTATAGAACTTAC 4525
OY 3961 ACCATGCGGATGTGTTTCCCGAGAAATGAGAACCAACCTGTCTCCGGGACCTTGAAGTC 4020
Db 4526 ACCATGCGGATGTGTTTCCCGAGAAATGAGAACCAACCTGTCTCTCCGGGACCTTGAAGTC 4585
OY 4021 CCGGCTTACCGGACAGAGCGTGTGAGAAAGGCTTGAAGCTTTGTGCCAGCTCATCAAC 4080
Db 4586 CCGGCTTACCGGACAGAGCGTGTGAGAAAGGCTTGAAGCTTTGTGCCAGCTCATCAAC 4645
OY 4081 AACAGATGTTCCTGTCTCTTCAATCCGAGAGCTTGAAGCTCCAGCGAGCTTCCGATG 4140
Db 4646 AACAGATGTTCCTGTCTCTTCAATCCGAGAGCTTGAAGCTCCAGCGAGCTTCCGATG 4705
OY 4141 CGCGACCGTGGCAAGTGGGCTCACTCATATGACCGTGTGACAGAGCAAGCTGAGTAC 4200
Db 4706 CGCGACCGTGGCAAGTGGGCTCACTCATATGACCGTGTGACAGAGCAAGCTGAGTAC 4765
OY 4201 GCGACTGATGTGTAAAGCATGTGTGGCCGACCTCATTTGACAAAGAACTGGAAGGCAAG 4260
Db 4766 GCGACTGATGTGTAAAGCATGTGTGGCCGACCTCATTTGACAAAGAACTGGAAGGCAAG 4825
OY 4261 AACCAACCTTAAAGCTCTGCTCAGAGAGCTGAGTCAAGTGTGAGAGATGTGCAAT 4320

Db 4826 AACCACTAAGTCTGCTGCTGAGAGAGCTGATGATGAGTGGCTGAGAAATGCTGACCAAT 4885
 QY 4321 TGGTTTACTTTCCTCCTCTACAGAGTTCCTCAAGGAGTGTGCTGGGAGAGCCCTCTTCTCC 4380
 QY 4886 TGGTTTACTTTCCTCCTCTACAGAGTTCCTCAAGGAGTGTGCTGGGAGAGCCCTCTTCTCC 4945
 QY 4381 CTGTTCTGTGCTTCAAGCAGAGATGAGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 4440
 Db 4946 CTGTTCTGTGCTTCAAGCAGAGATGAGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 5005
 QY 4441 GGGCCCTACTCCTTGAAGAGAGCAAGCTCATCCGCGCAGAGATTTGACTTCAAAACCTG 4500
 Db 5006 GCGCGCTACTCCTTGAAGAGAGCAAGCTCATCCGCGCAGAGATTTGACTTCAAAACCTG 5065
 QY 4501 GTCTGAGCTGTGTGAGCCAGAGCAATGCAACAGCCCGAGGTCCTCCAGTAAAGATCTTC 4560
 Db 5066 GTCTGAGCTGTGTGAGCCAGAGCAATGCAACAGCCCGAGGTCCTCCAGTAAAGATCTTC 5125
 QY 4561 AACTGTGACACATCACTCAAGTCAAGAGAAAGATTTGATGATGATGATGATGATGATGATG 4620
 Db 5126 AACTGTGACACATCACTCAAGTCAAGAGAAAGATTTGATGATGATGATGATGATGATGATG 5185
 QY 4621 CCTTGCTCCACCGGCGCCAAAGCTGAGATGATGATGATGATGATGATGATGATGATGATG 4680
 Db 5186 CCTTGCTCCACCGGCGCCAAAGCTGAGATGATGATGATGATGATGATGATGATGATGATG 5245
 QY 4681 GCAAGAGATATCTTGAGAGATGAAGACATCACCAAGATTTGAGATTTGAGAGGGA 4740
 Db 5246 GCAAGAGATATCTTGAGAGATGAAGACATCACCAAGATTTGAGATTTGAGAGGGA 5305
 QY 4741 CTGAACACTGAGCCCACTACAGAGTGCAGATGCTCCGATGCTGATGATGATGATGATGATG 4800
 Db 5306 CTGAACACTGAGCCCACTACAGAGTGCAGATGCTCCGATGCTGATGATGATGATGATGATG 5365
 QY 4801 CAGGTGACAGCTTATTAAGCAGATGAACAACTCCAGCTTCCAGAGCTTCAAGCAAGTAA 4860
 Db 5366 CAGGTGACAGCTTATTAAGCAGATGAACAACTCCAGCTTCCAGAGCTTCAAGCAAGTAA 5425
 QY 4861 TATGAAAACATGATCCGATGACAGGAGAGCCCGAGAGGCTCCGCTCAAGCAAGTAA 4920
 Db 5426 TATGAAAACATGATCCGATGACAGGAGAGCCCGAGAGGCTCCGCTCAAGCAAGTAA 5485
 QY 4921 ATCACTCTGACCTGAGAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGA 4980
 Db 5486 ATCACTCTGACCTGAGAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGA 5545
 QY 4981 GAGAGCAAG 5040
 Db 5546 GAGAGCAAG 5605
 QY 5041 CTCTCTGAGCACTAAGGAGCACTGAGAAAGTGTGATGATGATGATGATGATGATGATGATG 5100
 Db 5606 CTCTCTGAGCACTAAGGAGCACTGAGAAAGTGTGATGATGATGATGATGATGATGATGATG 5665
 QY 5101 AGAAGGAGCAAGCTGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 5160
 Db 5666 AGAAGGAGCAAGCTGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 5725
 QY 5161 GATGAGCAGGCTGATTAACATGATGATGATGATGATGATGATGATGATGATGATGATGATG 5220
 Db 5726 GATGAGCAGGCTGATTAACATGATGATGATGATGATGATGATGATGATGATGATGATGATG 5785
 QY 5221 AATGCTGCTGCTGAGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTG 5280
 Db 5786 AATGCTGCTGCTGAGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTG 5845
 QY 5281 ATCATATAGAAAGAGATCAAGAGAGCTGCTCTCTGATGATGATGATGATGATGATGATGATG 5340
 Db 5846 ATCATATAGAAAGAGATCAAGAGAGCTGCTCTCTGATGATGATGATGATGATGATGATGATG 5905
 QY 5341 TCTTGCTCAAGTCAAG 5400
 Db 5906 TCTTGCTCAAGTCAAG 5965

QY 5401 GCCAAGAGCATCCCGAGCTACAGAGATTTGGTGGAGAGTATTACTCAAGATAGGAGAG 5460
 Db 5966 GCCAAGAGCATCCCGAGCTACAGAGATTTGGTGGAGAGTATTACTCAAGATAGGAGAG 6025
 QY 5461 ATGCCAGGCATCAG 5520
 Db 6026 ATGCCAGGCATCAG 6085
 QY 5521 ATGATGAGTTCAACACATGATGATGATGATGATGATGATGATGATGATGATGATGATG 5580
 Db 6086 ATGATGAGTTCAACACATGATGATGATGATGATGATGATGATGATGATGATGATGATG 6145
 QY 5581 AGGAGAGAGATCCTTGAG 5640
 Db 6146 AGGAGAGAGATCCTTGAG 6205
 QY 5641 TACAACTAGAACAAAGTCAATTAACCTCATGAGCTTACAGAGCTGA 5685
 Db 6206 TACAACTAGAACAAAGTCAATTAACCTCATGAGCTTACAGAGCTGA 6250

RESULT 3
 ABK70006
 ID ABK70006 standard; DNA; 2597 BP.
 XX
 AC ABK70006;
 XX
 DT 15-JUL-2002 (first entry)
 XX
 DE cDNA encoding human Pro peptide #46.
 XX
 KM Human; ss; gene; PRO; secreted protein; transmembrane protein;
 KM genetic disorder; tumour; cancer.
 XX
 OS Homo sapiens.
 XX
 PN W0200224888-A2.
 PD 28-MAR-2002.
 XX
 PF 29-AUG-2001; 2001WO-US027099.
 XX
 PR 01-SEP-2000; 2000US-0229896P.
 PR 05-SEP-2000; 2000US-0230621P.
 PR 22-SEP-2000; 2000US-0235147P.
 PR 10-NOV-2000; 2000WO-US030873.
 PR 12-JAN-2001; 2001US-0261878P.
 PR 16-JAN-2001; 2001US-0261910P.
 PR 16-JAN-2001; 2001US-0261939P.
 PR 16-JAN-2001; 2001US-0262150P.
 PR 25-JAN-2001; 2001US-0264385P.
 PR 02-FEB-2001; 2001US-0264421P.
 PR 09-FEB-2001; 2001WO-US066520.
 PR 28-FEB-2001; 2001WO-US067623P.
 PR 09-MAR-2001; 2001US-0274399P.
 PR 03-APR-2001; 2001US-0280982P.
 PR 04-APR-2001; 2001US-0282129P.
 PR 04-APR-2001; 2001US-0282198P.
 PR 09-MAY-2001; 2001US-0280889P.
 PR 25-MAY-2001; 2001WO-US017092.
 PR 01-JUN-2001; 2001WO-US017800.
 PR 20-JUN-2001; 2001WO-US019692.
 PR 29-JUN-2001; 2001WO-US021066.
 PR 09-JUL-2001; 2001WO-US021735.
 XX
 PA (GERTH) GENENTECH INC.
 XX
 PI Baker KP, Eaton DL, Filvaroff E, Goddard A, Grimaldi JC,
 PI Gurney AL, Smith V, Stephan J, Watanabe CK, Wood WT, Zhang Z,
 PI Fong S,
 XX
 DR WPI; 2002-362426/39.

DR P-PSDB; ABG34075.
XX
XX New PRO polypeptides and polynucleotides encoding the polypeptides.
PT useful in gene therapy, chromosome identification, tissue typing, or for
PT genetic analysis of individuals with genetic disorders.
XX
XX
XX

Claim 2; Fig 91; 218bp; English.

CC This invention relates to the cDNA and protein sequences of novel
CC secreted and transmembrane polypeptides PRO polypeptides. The invention
CC also comprises a method for producing the proteins of the invention by
CC recombinant means and antibodies specific for the protein of the
CC invention. The antibody may be used for detecting the PRO proteins of the
CC invention and may be used to modify their activity. polynucleotides may
CC be used as hybridisation probes for a cDNA library to isolate the full-
CC length PRO cDNA or to isolate other cDNAs, to construct hybridisation
CC probes for mapping the gene which encodes that PRO and for genetic
CC analysis of individuals with genetic disorders, in assays to identify
CC other proteins or molecules involved in binding reaction, to generate
CC transgenic animals or knock-out animals which in turn are useful in the
CC development and screening of therapeutically useful reagents, for
CC chromosome identification, and tissue typing. The PRO polypeptides are
CC useful in gene therapy, and as molecular weight markers for protein
CC electrophoresis purposes. The sequences may also be used to detect
CC overexpression on PRO polypeptides in cancerous tumours and for screening
CC for differentially expressed genes using microarray technology. The
CC present sequence represents a cDNA encoding a human PRO protein of the
CC invention
XX

XX Sequence 2597 BP; 688 A; 570 C; 642 G; 697 T; 0 U; 0 Other;

XX Query Match 13.9%; Score 793; DB 6; Length 2597;

XX Best Local Similarity 100.0%; Pred. No. 0;

XX Matches 793; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

XX 579 CGAGTATTTTCCGACATCTCCAGCCGAAACTGACCAAGAACTCTGAGCGGATGGCAT 638

XX 1 CGAGTATTTTCCGACATCTCCAGCCGAAACTGACCAAGAACTCTGAGCGGATGGCAT 60

XX 639 GTTCGCGTACGCTCTTCATATGATGATGCTGCGCTGATGATTAAGATCCCTTCGACAC 698

XX 61 GTTCGCGTACGCTCTTCATATGATGATGCTGCGCTGATGATTAAGATCCCTTCGACAC 120

XX 699 CTTCACCATCATCCCTGACTTTGATATCTACTATGCTATGCTTATGAGAGTGGCACTT 758

XX 121 CTTCACCATCATCCCTGACTTTGATATCTACTATGCTATGCTTATGAGAGTGGCACTT 180

XX 759 TGTCTACTTTTGAACCTTCGACCTGATGATGCTGCTCCACAGGCTCCACCAAGA 818

XX 181 TGTCTACTTTTGAACCTTCGACCTGATGATGCTGCTCCACAGGCTCCACCAAGA 240

XX 819 GCAGGTGATATACCAAGAGCTGCTGAGAGCTTTGCAAGAGACAGCCCTTCACTCTTA 878

XX 241 GCAGGTGATATACCAAGAGCTGCTGAGAGCTTTGCAAGAGACAGCCCTTCACTCTTA 300

XX 879 TGTAGAGGTGCGCCATTGGCTGTGAGCGAGTGGGCTGAGTACCGCTGCTGAGGCTGC 938

XX 301 TGTAGAGGTGCGCCATTGGCTGTGAGCGAGTGGGCTGAGTACCGCTGCTGAGGCTGC 360

XX 939 CTACCTGTCCAAAGCGGGGCGCTGCTGAGAGACCTTGGATCCATCAATGATGA 998

XX 361 CTACCTGTCCAAAGCGGGGCGCTGCTGAGAGACCTTGGATCCATCAATGATGA 420

XX 999 CCTGCTCTTCAACCGCTCTTCTCCAGGCGCAGAAACGGAATAATCCTGTGATGATG 1058

XX 421 CCTGCTCTTCAACCGCTCTTCTCCAGGCGCAGAAACGGAATAATCCTGTGATGATG 480

XX 1059 GGCCTGTGCACTTCTCATCTTGAAGACAGATTAATGACCGATTAAGAGAGGCGCTGACATC 1118

XX 481 GGCCTGTGCACTTCTCATCTTGAAGACAGATTAATGACCGATTAAGAGAGGCGCTGACATC 540

XX 1119 TTGTTACCGGGGCGAGGCGACGCTGAGCTGAGCTGAGCTGAGTGAAGACATCCCTTG 1178

Db 541 TTGTTACCGGGGCGAGGCGACGCTGAGCTGAGCTGAGCTGAGTGAAGACATCCCTTG 600

Qy 1179 CAGAGTGCCTCTTAAACATTAAGATTAATCTGTGGCGCTTGACATGATGTCCTCT 1238

Db 601 CAGAGTGCCTCTTAAACATTAAGATTAATCTGTGGCGCTTGACATGATGTCCTCT 660

Qy 1239 GCGAGTGTCCGACATGCTGTGGATATCCCTCTTCAAGAGAGAGAGGACCGCATGAC 1298

Db 661 GCGAGTGTCCGACATGCTGTGGATATCCCTCTTCAAGAGAGAGAGGACCGCATGAC 720

Qy 1299 GTCTGTGATCGCATATGCTTACAGAACCACTCTGCTTGTGGCGACCAAAAGTG 1358

Db 721 GTCTGTGATCGCATATGCTTACAGAACCACTCTGCTTGTGGCGACCAAAAGTG 780

Qy 1359 CAAGCTGAGAGAG 1371

Db 781 CAAGCTGAGAGAG 793

RESULT 4

ADA01359

ID ADA01359 standard; cDNA; 2597 BP.

XX ADA01359;

AC ADA01359;

XX 06-NOV-2003 (first entry)

DT Human PRO polynucleotide #46.

XX Human; gene; ss; PRO; secreted polypeptide; transmembrane polypeptide;

XX tumour necrosis factor-alpha; TNF-alpha; blood; chondrocyte cell; tumour;

XX adrenal; lung; colon; breast; prostate; rectum; cervix; liver; cancer;

XX microvascular endothelial cell; endothelial cell tube formation;

XX sports-related joint problem; articular cartilage defect; osteoarthritis;

XX rheumatoid arthritis; osteopathic; anti-rheumatic; antiarthritic.

XX Homo sapiens.

XX US2003068779-A1.

XX 10-APR-2003.

XX 16-SEP-2002; 2002US-00245107.

XX 09-MAY-2001; 2001US-0230589P.

XX 29-AUG-2001; 2001WO-US027099.

XX 18-JUL-2002; 2002US-00197942.

XX (GENENTECH INC.

XX Baker KP, Baton DL, Filvaroff E, Goddard A, Grimaldi JC,

XX Gurney AL, Smith V, Stephan JF, Watanabe CK, Wood WT, Zhang Z,

XX Fong S;

XX MPI: 2003-625484/59.

XX P-PSDB; ADA01360.

XX Novel isolated PRO1313, PRO20080 or PRO21383 polypeptide useful for

XX stimulating proliferation of human microvascular endothelial cells, and

XX PRO6018 polypeptide useful for stimulating proliferation of chondrocyte

XX cells.

XX Claim 2; Fig 91; 307bp; English.

XX The invention relates to isolated human PRO polypeptides (secreted and

XX transmembrane polypeptides) and the polynucleotides encoding them. The

XX invention also relates to an antibody which specifically binds to a PRO

XX polypeptide, a method for stimulating the release of tumour necrosis

XX factor-alpha (TNF-alpha) from human blood, a method for stimulating the

XX proliferation or differentiation of chondrocyte cells and a method for

XX detecting the presence of a tumour in a mammal (e.g. adrenal, lung,

XX colon, breast, prostate, rectal, cervical and liver tumours). The

XX polynucleotides are useful in molecular biology, including uses as

hybridisation probes, in chromosome and gene mapping, in generating antisense RNA and DNA and in gene therapy. The polynucleotides may also be used in preparing PRO polypeptides by recombinant techniques and in generating either transgenic animals or knock-out animals which are useful in the development and screening of therapeutically useful reagents. The PRO polypeptides or antibodies are used in preparing a medicament for treating a condition responsive to the polypeptides or antibodies, such as tumours, for stimulating and inhibiting proliferation of human microvascular endothelial cells, for inducing endothelial cell tube formation and for treating sports-related problems, articular cartilage defects, osteoarthritis and rheumatoid arthritis. This sequence represents a human PRO polynucleotide of the invention.

Sequence 2597 BP; 688 A; 570 C; 642 G; 697 T; 0 U; 0 Other;

Query Match 13.9%; Score 793; DB 8; Length 2597;

Best Local Similarity 100.0%; Pred. No. 0;

Matches 793; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

579 CGAGTATTTTCCCACTCTCCAGCCCGAATCTGACCAAGAACTCTGAGCGGATGGCAT 638
 1 CGAGTATTTTCCCACTCTCCAGCCCGAATCTGACCAAGAACTCTGAGCGGATGGCAT 60
 639 GTTCCGTAAGCTCTCCATGATGATGATGATGATGATGATGATGATGATGATGATGAT 698
 61 GTTCCGTAAGCTCTCCATGATGATGATGATGATGATGATGATGATGATGATGATGAT 120
 699 GTTCCGTAAGCTCTCCATGATGATGATGATGATGATGATGATGATGATGATGATGAT 758
 121 GTTCCGTAAGCTCTCCATGATGATGATGATGATGATGATGATGATGATGATGATGAT 180
 759 GTTCCGTAAGCTCTCCATGATGATGATGATGATGATGATGATGATGATGATGATGAT 818
 181 GTTCCGTAAGCTCTCCATGATGATGATGATGATGATGATGATGATGATGATGATGAT 240
 819 GTTCCGTAAGCTCTCCATGATGATGATGATGATGATGATGATGATGATGATGATGAT 878
 241 GTTCCGTAAGCTCTCCATGATGATGATGATGATGATGATGATGATGATGATGATGAT 300
 879 GTTCCGTAAGCTCTCCATGATGATGATGATGATGATGATGATGATGATGATGATGAT 938
 301 GTTCCGTAAGCTCTCCATGATGATGATGATGATGATGATGATGATGATGATGATGAT 360
 939 GTTCCGTAAGCTCTCCATGATGATGATGATGATGATGATGATGATGATGATGATGAT 998
 361 GTTCCGTAAGCTCTCCATGATGATGATGATGATGATGATGATGATGATGATGATGAT 420
 999 GTTCCGTAAGCTCTCCATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1058
 421 GTTCCGTAAGCTCTCCATGATGATGATGATGATGATGATGATGATGATGATGATGAT 480
 1059 GTTCCGTAAGCTCTCCATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1118
 481 GTTCCGTAAGCTCTCCATGATGATGATGATGATGATGATGATGATGATGATGATGAT 540
 1119 GTTCCGTAAGCTCTCCATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1178
 541 GTTCCGTAAGCTCTCCATGATGATGATGATGATGATGATGATGATGATGATGATGAT 600
 1179 GTTCCGTAAGCTCTCCATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1238
 601 GTTCCGTAAGCTCTCCATGATGATGATGATGATGATGATGATGATGATGATGATGAT 660
 1239 GTTCCGTAAGCTCTCCATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1298
 661 GTTCCGTAAGCTCTCCATGATGATGATGATGATGATGATGATGATGATGATGATGAT 720
 1299 GTTCCGTAAGCTCTCCATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1358
 721 GTTCCGTAAGCTCTCCATGATGATGATGATGATGATGATGATGATGATGATGATGAT 780
 1359 GTTCCGTAAGCTCTCCATGATGATGATGATGATGATGATGATGATGATGATGATGAT

Db 781 CAACTGAGAGAG 793

RESULT 5

ADA3788

ADA3788 standard; cDNA; 2597 BP.

ADA3788;

20-NOV-2003 (first entry)

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

65; gene; human; PRO; secreted protein; transmembrane protein; endothelial cell tube formation; chondrocyte cell differentiation; microvascular endothelial cell; tumour; lung tumour; colon tumour; breast tumour; prostate tumour; rectal tumour; kidney tumour; liver tumour; cytostatic; vaccine.

Homo sapiens.

US2003064474-A1.

03-APR-2003.

16-SHP-2002; 2002US-00245859.

29-AUG-2001; 2001WO-US027099.

18-JUL-2002; 2002US-00197942.

(GENTH) GENENTECH INC.

Baker KP, Eaton DL, Filvaroff E, Goddard A, Grimaldi JC, Gurney AL, Smith V, Stephan JP, Watanabe CK, Wood WL, Zhang Z, Fong S;

WPI; 2003-605867/57.

P-PSDB; ADA3789.

New isolated nucleic acid encoding a PRO polypeptide, e.g. PRO20080 or PRO21383, useful in molecular biology, chromosome and gene mapping, in generating antisense RNA and DNA, and in gene therapy.

Claim 2; Fig 91; 308pp; English.

The invention relates to an isolated secreted/transmembrane (PRO) polypeptide, having at least 80% sequence identity to a sequence selected from any one of the 57 amino acid sequences given in specification, or to a sequence encoded by a nucleic acid molecule selected from any one of the nucleic acids deposited under any of the ATCC accession numbers given in specification, or a sequence having at least 80% identity to PRO lacking its associated signal peptide, an extracellular domain of PRO with or without its associated signal peptide. Also included are vectors, transformed host cells, anti-PRO antibodies, the nucleic acids encoding PRO, PRO fusion proteins, inducing endothelial cell tube formation (by administering PRO281, PRO1560, PRO189, PRO499, PRO6308, PRO6000, PRO10275, PRO2107, PRO20933 or PRO34274 polypeptide or its agonist) and an oligonucleotide probe derived from any one of the above nucleotide sequences. PRO6018 polypeptide is useful for stimulating the proliferation or differentiation of chondrocyte cells. PRO1313, PRO20080 and PRO21383 polypeptides are useful for stimulating the proliferation of human microvascular endothelial cells. PRO6071, PRO487 and PRO6006 polypeptides are useful for inhibiting the proliferation of human microvascular endothelial cells. PRO polypeptides are useful for detecting the presence of tumour in a mammal, including tumours of lung, colon, breast, prostate, rectal, kidney and liver. PRO281, PRO1560, PRO189, PRO499, PRO6308, PRO6000, PRO10275, PRO2107, PRO20933 and PRO34274 polypeptides are useful for inducing endothelial cell tube formation. PRO or the antibody are useful in the preparation of a medicament for treating a condition responsive to PRO polypeptide. The oligonucleotide probes are useful for isolating genomic and cDNA nucleotide sequences, for measuring or detecting the expression of an associated gene, and as antisense probes. PRO nucleic acid is useful as a

CC hybridisation probe, in chromosome and gene mapping, in the generation of
CC antisense RNA and DNA, and for the preparation PRO polypeptides. PRO and
CC PRO nucleic acid are useful as therapeutic agents, e.g. vaccines. The
CC present sequence encodes a PRO protein.

XX Sequence 2597 BP; 688 A; 570 C; 642 G; 697 T; 0 U; 0 Other;

Query Match 13.9%; Score 793; DB 8; Length 2597;

Best Local Similarity 100.0%; Pred. No. 0;
Matches 793; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

QY 579 CGAGTATTTTCCACCATCTCCAGCCGGAACTCAACCAAGACTCTTGAAGCGGATGGCAT 638
DB 1 CGAGTATTTTCCACCATCTCCAGCCGGAACTCAACCAAGACTCTTGAAGCGGATGGCAT 60
QY 639 GTTCCGCTACGCTTCATGATGAGTTCGTGGCTTCATGATTAAATCCCTTCGACAC 698
DB 61 GTTCCGCTACGCTTCATGATGAGTTCGTGGCTTCATGATTAAATCCCTTCGACAC 120
QY 699 CTTCACCATTCATCCCTGACTTGTATCTATCTATGCTATGCTTTAGCAGTGGCACTT 758
DB 121 CTTCACCATTCATCCCTGACTTGTATCTATCTATGCTATGCTTTAGCAGTGGCACTT 180
QY 759 TGTCTACTTTTGAACCTCCAGCTGAGATGATGCTGCTCCACAGGCTCCACCAAGAGA 818
DB 181 TGTCTACTTTTGAACCTCCAGCTGAGATGATGCTGCTCCACAGGCTCCACCAAGAGA 240
QY 819 GCAGGTGATATCATCTCAAGCTCTGAGGCTTTGCAAGAGACACAGCTTCAACTCTTA 878
DB 241 GCAGGTGATATCATCTCAAGCTCTGAGGCTTTGCAAGAGACACAGCTTCAACTCTTA 300
QY 879 TGTAGAGTGGCCCATTTGCTGTGAGCGAGTGGAGTACCGCTGCGACAGGCTGC 938
DB 301 TGTAGAGTGGCCCATTTGCTGTGAGCGAGTGGAGTACCGCTGCGACAGGCTGC 360
QY 939 CTACCTGTCCAAAGCGGGGCGGTGCTTGCGAGNACCTTGTGATCCATTCATGATGA 998
DB 361 CTACCTGTCCAAAGCGGGGCGGTGCTTGCGAGNACCTTGTGATCCATTCATGATGA 420
QY 999 CCTGCTTTTACCGCTTCTTCCAGGGGCAAGCGGAAATGAAATCCCTGATGATGC 1058
DB 421 CCTGCTTTTACCGCTTCTTCCAGGGGCAAGCGGAAATGAAATCCCTGATGATGC 480
QY 1059 GGCCTGTGCACTTCTTCAAGAGCAGATTAATGACCGCATTAAGAGCGGCTGCAATC 1118
DB 481 GGCCTGTGCACTTCTTCAAGAGCAGATTAATGACCGCATTAAGAGCGGCTGCAATC 540
QY 1119 TTGTATCCGGGGCGAGGGGCAAGCTGGCTGGCTCAAGGTGAAGAGCATCCCTG 1178
DB 541 TTGTATCCGGGGCGAGGGGCAAGCTGGCTGGCTCAAGGTGAAGAGCATCCCTG 600
QY 1179 CAGCAGTGGGCTTAAACATGATGATGATGATGATGATGATGATGATGATGATGATG 1238
DB 601 CAGCAGTGGGCTTAAACATGATGATGATGATGATGATGATGATGATGATGATGATG 660
QY 1239 GGGAGTGTCCGACATGCTGCTGGAATTCCTTTCACGAGAGACAGGAGCCGATGAC 1298
DB 661 GGGAGTGTCCGACATGCTGCTGGAATTCCTTTCACGAGAGACAGGAGCCGATGAC 720
QY 1299 GGTCTGTCATTCGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1358
DB 721 GGTCTGTCATTCGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 780
QY 1359 CAAAGTGAAGAG 1371
DB 781 CAAAGTGAAGAG 793

```

RESULT 6
ADA43556
ID ADA43556 standard; cDNA; 2597 BP.
XX
AC ADA43556;

```

XX 20-NOV-2003 (first entry)
XX Human cDNA encoding secreted/transmembrane polypeptide PRO34003.
DE ss; gene; human; PRO; secreted protein; transmembrane protein;
XX endotheial cell tube formation; chondrocyte cell differentiation;
XX microvascular endothelial cell; tumour; lung tumour; colon tumour;
XX breast tumour; prostate tumour; rectal tumour; kidney tumour;
XX liver tumour; cytosolic; vaccine.
XX Homo sapiens.
XX US2003073196-A1.
XX 17-APR-2003.
XX 18-SEP-2002; 2002US-00246210.
XX 04-APR-2001; 2001US-0282199P.
XX 29-APR-2001; 2001MO-US027039.
XX 18-JUL-2002; 2002US-00197942.
XX (GENT) GENENTECH INC.
XX Baker KP, Eaton DL, Filvaroff E, Goddard A, Grimaldi JC,
XX Gurney AL, Smith V, Stephan JP, Watanabe CK, Wood WI, Zhang Z;
XX Fong S;
XX WPI; 2003-743814/70.
XX P-PDB; ADA43557.
XX DR New secreted and transmembrane PRO polypeptides, e.g. PRO20080 or
XX PRO21383 useful for stimulating the proliferation or differentiation of
XX chondrocyte cells and detecting the presence of a tumor in a mammal.
XX Claim 2; Fig 91; 307pp; English.
XX The invention relates to an isolated secreted/transmembrane (PRO)
XX polypeptide, having at least 80% sequence identity to a sequence selected
XX from any one of the 57 amino acid sequences given in specification, or to
XX a sequence encoded by a nucleic acid molecule selected from any one of
XX the nucleic acids deposited under any of the ATCC accession numbers given
XX in specification, or a sequence having at least 80% identity to PRO
XX lacking its associated signal peptide. Also included are vectors,
XX with or without its associated signal peptide. Also included are vectors,
XX transformed host cells, anti-PRO antibodies, the nucleic acids encoding
XX PRO, PRO fusion proteins, inducing endothelial cell tube formation (by
XX administering PRO281, PRO1560, PRO189, PRO4499, PRO6308, PRO6000,
XX PRO10275, PRO21207, PRO20933 or PRO34274 polypeptide or its agonist) and
XX an oligonucleotide probe derived from any one of the above nucleotide
XX sequences. PRO6018 polypeptide is useful for stimulating the
XX proliferation or differentiation of chondrocyte cells. PRO1313, PRO20080
XX and PRO21383 polypeptides are useful for stimulating the proliferation of
XX human microvascular endothelial cells. PRO6071, PRO4487 and PRO6006
XX polypeptides are useful for inhibiting the proliferation of human
XX microvascular endothelial cells. PRO polypeptides are useful for
XX detecting the presence of tumour in a mammal, including tumours of lung,
XX colon, breast, prostate, rectal, kidney and liver. PRO281, PRO1560,
XX PRO189, PRO4499, PRO6308, PRO10275, PRO21207, PRO20933 and
XX PRO34274 polypeptides are useful for inducing endothelial cell tube
XX formation. PRO or the antibody are useful in the preparation of a
XX medicament for treating a condition responsive to PRO polypeptide. The
XX oligonucleotide probes are useful for isolating genomic and cDNA
XX associated sequences, for measuring or detecting the expression of an
XX associated gene, and as antisense probes. PRO nucleic acid is useful as a
XX hybridisation probe, in chromosome and gene mapping, in the generation of
XX antisense RNA and DNA, and for the preparation PRO polypeptides. PRO and
XX PRO nucleic acid are useful as therapeutic agents, e.g. vaccines. The
XX present sequence encodes a PRO protein.
XX Sequence 2597 BP; 688 A; 570 C; 642 G; 697 T; 0 U; 0 Other;

```

Query Match 13.9%; Score 793; DB 8; Length 2597;
 Best Local Similarity 100.0%; Pred. No. 0;
 Matches 793; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

QY 579 CGAGTATTTTCCACCATCTCCAGCCGGAACCTGACCAAGAACTCTGAGCGGATGCGAT 638
DB 1 CGAGTATTTTCCACCATCTCCAGCCGGAACCTGACCAAGAACTCTGAGCGGATGCGAT 60
QY 639 GTTCCGACGCTCTTCCATGATGATGAGTGGCCCGATGATTAAGATCCCTTGGACAC 698
DB 61 GTTCCGACGCTCTTCCATGATGATGAGTGGCCCGATGATTAAGATCCCTTGGACAC 120
QY 699 CTTCACCATCATCCCTGACTTGTATATCTACTATGTCTATGCTTTTAAAGATGCGAATT 758
DB 121 CTTCACCATCATCCCTGACTTGTATATCTACTATGTCTATGCTTTTAAAGATGCGAATT 180
QY 759 TGTCTACTTTTGAACCTCCACCTGAGATGCTGCTCCACAGGCTCCACCAAGAA 818
DB 181 TGTCTACTTTTGAACCTCCACCTGAGATGCTGCTCCACAGGCTCCACCAAGAA 240
QY 819 GCGAGTGTATACATCCAGCTGCTGAGGCTTTGCAAGAGACACAGCTTCAACTCTTA 878
DB 241 GCGAGTGTATACATCCAGCTGCTGAGGCTTTGCAAGAGACACAGCTTCAACTCTTA 300
QY 879 TGTAGAGTGGCCATTTGGCTGTGAGCGCAGTGGGCTGAGATACCGCTCCGACAGGCTGC 938
DB 301 TGTAGAGTGGCCATTTGGCTGTGAGCGCAGTGGGCTGAGATACCGCTCCGACAGGCTGC 360
QY 939 CTACCTGTCCAAAGCGGGGCGCTGCTTGGCAGAGCCCTTGGAGTCCATCCAGATGATGA 998
DB 361 CTACCTGTCCAAAGCGGGGCGCTGCTTGGCAGAGCCCTTGGAGTCCATCCAGATGATGA 420
QY 999 CCTGCTCTTCAACCGCTTTCTCCAGGGCCAGAAAGGGGAAATCCCTGATGATGATC 1058
DB 421 CCTGCTCTTCAACCGCTTTCTCCAGGGCCAGAAAGGGGAAATCCCTGATGATGATC 480
QY 1059 GGGCCCTGTGCATCTTCACTTGAAGAGATTAATGACCGCATTAAGAGACGGCTGCAGTC 1118
DB 481 GGGCCCTGTGCATCTTCACTTGAAGAGATTAATGACCGCATTAAGAGACGGCTGCAGTC 540
QY 1119 TTGTACCGGGGCGAGGGGACGCTGAGACCTGCGCTGCAAGGTGAAGAGATCCCTG 1178
DB 541 TTGTACCGGGGCGAGGGGACGCTGAGACCTGCGCTGCAAGGTGAAGAGATCCCTG 600
QY 1179 CAGCAATGTGGCTCTTAACCATTAAGCATTAATCTGTGCGCTGCAATGATGCTCCCT 1238
DB 601 CAGCAATGTGGCTCTTAACCATTAAGCATTAATCTGTGCGCTGCAATGATGCTCCCT 660
QY 1239 GGGAGTGTCCGACATGCTGTGGAATTCCTGCTTCCAGAGAGACAGGACCGCATGAC 1298
DB 661 GGGAGTGTCCGACATGCTGTGGAATTCCTGCTTCCAGAGAGACAGGACCGCATGAC 720
QY 1299 GTCTGCATGCGCATATGCTTACAGAAACCACTCTGCGCTTGTGCGGACCAAAAGTGG 1358
DB 721 GTCTGCATGCGCATATGCTTACAGAAACCACTCTCTGCGCTTGTGCGGACCAAAAGTGG 780
QY 1359 CAAGCTGAAGAG 1371
DB 781 CAAGCTGAAGAG 793

```

RESULT 7

ADA01231 standard; cDNA; 2597 BP.

ADA01231;

06-NOV-2003 (first entry)

Human PRO polynucleotide #46.

Human; gene; ss; PRO; secreted polypeptide; transmembrane polypeptide;

tumour necrosis factor-alpha; TNF-alpha; blood; chondrocyte cell; tumour;

KM Cancer; lung; colon; breast; prostate; rectum; kidney; liver;
 KW microvascular endothelial cell; endothelial cell tube formation.
 XX Homo sapiens.

US2003068782-A1.

10-APR-2003.

16-SEP-2002; 2002US-00245851.

27-APR-1999; 99US-0131271P.

28-OCT-1999; 99US-0162506P.

02-DEC-1999; 99WO-US028551.

29-AUG-2001; 2001WO-US027099.

18-JUL-2002; 2002US-00197942.

(GENTH) GENENTECH INC.

Baker KP, Eaton DL, Filvaroff E, Goddard A, Grimaldi JC,
 Guirney AL, Smith V, Stephan JP, Watanabe CX, Wood WI, Zhang Z,
 Fong S;

WPI; 2003-625487/59.

P-PSDB; ADA01232.

Novel isolated PRO polypeptides e.g. PRO281 and PRO1560, useful in the
 preparation of a medicament for treating a condition responsive to PRO
 polypeptide, and as therapeutic agents e.g. vaccines.

Claim 2; Fig 91; 308bp; English.

The invention relates to isolated human PRO polypeptides (secreted and
 transmembrane polypeptides) and the polynucleotides encoding them. The
 invention also relates to an antibody which specifically binds to a PRO
 polypeptide, a method for stimulating the release of tumour necrosis
 factor-alpha (TNF-alpha) from human blood, a method for stimulating the
 proliferation or differentiation of chondrocyte cells and a method for
 detecting the presence of a tumour in a mammal (e.g. lung, colon, breast,
 prostate, rectal, kidney and liver tumours). The polynucleotides are
 useful in molecular biology, including uses as hybridisation probes, in
 chromosome and gene mapping, in generating antisense RNA and DNA and in
 gene therapy. The polynucleotides may also be used in preparing PRO
 polypeptides by recombinant techniques and in generating either
 transgenic animals or knock-out animals which are useful in the
 development and screening of therapeutically useful reagents. The PRO
 polypeptides or antibodies are used in preparing a medicament for
 treating a condition responsive to the polypeptides or antibodies, such
 as tumours, for stimulating and inhibiting proliferation of human
 microvascular endothelial cells and for inducing endothelial cell tube
 formation. This sequence represents a human PRO polynucleotide of the
 invention.

Sequence 2597 BP; 688 A; 570 C; 642 G; 697 T; 0 U; 0 Other;

Query Match 13.9%; Score 793; DB 8; Length 2597;

Best Local Similarity 100.0%; Pred. No. 0;
 Matches 793; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

QY 579 CGAGTATTTTCCACCATCTCCAGCCGGAACCTGACCAAGAACTCTGAGCGGATGCGAT 638
DB 1 CGAGTATTTTCCACCATCTCCAGCCGGAACCTGACCAAGAACTCTGAGCGGATGCGAT 60
QY 639 GTTCCGACGCTCTTCCATGATGATGAGTGGCCCGATGATTAAGATCCCTTGGACAC 698
DB 61 GTTCCGACGCTCTTCCATGATGATGAGTGGCCCGATGATTAAGATCCCTTGGACAC 120
QY 699 CTTCACCATCATCCCTGACTTGTATATCTACTATGTCTATGCTTTTAAAGATGCGAATT 758
DB 121 CTTCACCATCATCCCTGACTTGTATATCTACTATGTCTATGCTTTTAAAGATGCGAATT 180
QY 759 TGTCTACTTTTGAACCTCCACCTGAGATGCTGCTCCACAGGCTCCACCAAGAA 818

```

Db 181 TGTCTACTTTTGAACCTCCAACTGAGATGATGTGTCTCCACCAAGGCTCCACCAAGCA 240
 QY 819 GCAAGTGTATATACATCCAGAGCTGTGAGGTTTGGCAAGAGACACAGCTTCAATCCTTA 878
 Db 241 GCAAGTGTATATACATCCAGAGCTGTGAGGTTTGGCAAGAGACACAGCTTCAATCCTTA 300
 QY 879 TGTAGAGTGTCCATTTGCTGTAGAGGCAAGGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTG 938
 Db 301 TGTAGAGTGTCCATTTGCTGTAGAGGCAAGGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTG 360
 QY 939 CTACTGTGTCCAAAGCGGAGGCGCTGTGAGGAGACCTTGTGAGTGTGAGTGTGAGTGTGAGTGTG 998
 Db 361 CTACTGTGTCCAAAGCGGAGGCGCTGTGAGGAGACCTTGTGAGTGTGAGTGTGAGTGTGAGTGTG 420
 QY 999 CCGTGTCTTCAAGCGTGTGTCTCCAAAGGCGCAAGAGCGGAAATGAAATCCCTGTGATGAGTTC 1058
 Db 421 CCGTGTCTTCAAGCGTGTGTCTCCAAAGGCGCAAGAGCGGAAATGAAATCCCTGTGATGAGTTC 480
 QY 1059 GAGCCCTGTGATCTTCAATCTTGAAGCAGATTAATGACCGCATTTAAGAGCGGCTGTGAGTTC 1118
 Db 481 GAGCCCTGTGATCTTCAATCTTGAAGCAGATTAATGACCGCATTTAAGAGCGGCTGTGAGTTC 540
 QY 1119 TTTTTCACCGGAGGAGGAGGACCGTGTGACCTGTGCTGTGAGTGTGAGTGTGAGTGTGAGTGTG 1178
 Db 541 TTTTTCACCGGAGGAGGAGGACCGTGTGACCTGTGCTGTGAGTGTGAGTGTGAGTGTGAGTGTG 600
 QY 1179 CAGCAGTGTGCTCTTCAATCTTGAAGCAGATTAATGACCTGTGCTGTGAGTGTGAGTGTGAGTGTG 1238
 Db 601 CAGCAGTGTGCTCTTCAATCTTGAAGCAGATTAATGACCTGTGCTGTGAGTGTGAGTGTGAGTGTG 660
 QY 1239 GGGAGTGTCCGATGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTG 1298
 Db 661 GGGAGTGTCCGATGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTG 720
 QY 1299 GTCTGTATGATGATGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTG 1358
 Db 721 GTCTGTATGATGATGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTG 780
 QY 1359 CAAAGCTGAGAGAG 1371
 Db 781 CAAAGCTGAGAGAG 793
 RESULT 8
 ID ADA01115 standard; cDNA; 2597 BP.
 XX ADA01115;
 XX 06-NOV-2003 (first entry)
 XX Human cDNA encoding secreted/transmembrane polypeptide PRO34003.
 DE ss; gene; human; PRO; secreted protein; transmembrane protein;
 XX endothelial cell tube formation; chondrocyte cell differentiation;
 KW microvascular endothelial cell; tumour; lung tumour; colon tumour;
 KW breast tumour; prostate tumour; rectal tumour; kidney tumour;
 KW liver tumour; cytostatic; vaccine.
 XX Homo sapiens.
 OS US2003068780-A1.
 XX PD 10-Apr-2003.
 XX PF 16-SEP-2002; 2002US-00245143.
 XX PR 02-AUG-2000; 2000US-0222695P.
 XX PR 20-JUN-2001; 2001WO-US019692.
 XX PR 29-AUG-2001; 2001WO-US027099.
 XX PR 18-JUL-2002; 2002US-00197942.
 XX PA (GETH) GENENTECH INC.

XX Baker KP, Baton DV, Filvaroff E, Goddard A, Grimaldi JC;
 PI Gurney AL, Smith V, Stephan JP, Watanabe CK, Wood WI, Zhang Z;
 PI Fong S;
 XX WP: 2003-625485/59.
 DR P-PSDB; ADA01116.
 XX Novel isolated PRO polypeptides e.g. PRO281 and PRO1560, useful in the
 PT preparation of a medicament for treating a condition responsive to PRO
 PT polypeptide, and as therapeutic agents e.g. vaccines.
 PS Claim 2; Fig 91; 307pp; English.
 XX The invention relates to an isolated secreted/transmembrane (PRO)
 CC polypeptide, having at least 80% sequence identity to a sequence selected
 CC from any one of the 57 amino acid sequences given in specification, or to
 CC a sequence encoded by a nucleic acid molecule selected from any one of
 CC the nucleic acids deposited under any of the ATCC accession numbers given
 CC in specification, or a sequence having at least 80% identity to PRO
 CC lacking its associated signal peptide, an extracellular domain of PRO
 CC with or without its associated signal peptide. Also included are vectors,
 CC transformed host cells, anti-PRO antibodies, the nucleic acids encoding
 CC PRO, PRO fusion proteins, inducing endothelial cell tube formation (by
 CC administering PRO281, PRO1560, PRO499, PRO6308, PRO6000,
 CC PRO10275, PRO21207, PRO20933 or PRO34274 polypeptide or its agonist) and
 CC an oligonucleotide probe derived from any one of the above nucleotide
 CC sequences. PRO6018 polypeptide is useful for stimulating the
 CC proliferation or differentiation of chondrocyte cells. PRO313, PRO20080
 CC and PRO21383 polypeptides are useful for stimulating the proliferation of
 CC human microvascular endothelial cells. PRO6071, PRO4487 and PRO6006
 CC polypeptides are useful for inhibiting the proliferation of human
 CC microvascular endothelial cells. PRO polypeptides are useful for
 CC detecting the presence of tumour in a mammal, including tumours of lung,
 CC colon, breast, prostate, rectal, kidney and liver. PRO281, PRO1560,
 CC PRO189, PRO4499, PRO6308, PRO6000, PRO10275, PRO20933 and
 CC PRO34274 polypeptides are useful for inducing endothelial cell tube
 CC formation. PRO or the antibody are useful in the preparation of a
 CC medicament for treating a condition responsive to PRO polypeptide. The
 CC oligonucleotide probes are useful for isolating genomic and cDNA
 CC nucleotide sequences, for measuring or detecting the expression of an
 CC associated gene, and as antisense probes. PRO nucleic acid is useful as a
 CC hybridisation probe, in chromosome and gene mapping, in the generation of
 CC antisense RNA and DNA, and for the preparation PRO polypeptides. PRO and
 CC PRO nucleic acid are useful as therapeutic agents, e.g. vaccines. The
 CC present sequence encodes a PRO protein.
 XX Sequence 2597 BP; 688 A; 570 C; 642 G; 697 T; 0 U; 0 Other;
 SQ
 Query Match 13.9%; Score 793; DB 8; Length 2597;
 Best Local Similarity 100.0%; Pred. No. 0;
 Matches 793; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 579 GAGATTTTCCCAACATCTCCAGCGGAACTGACCAAGACCTGTGAGGCGATGGCAT 638
 Db 1 CAGATTTTCCCAACATCTCCAGCGGAACTGACCAAGACCTGTGAGGCGATGGCAT 60
 QY 639 GTTCCGATGCTTTCATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 698
 Db 61 GTTCCGATGCTTTCATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 120
 QY 699 CTTACCAATATCCCTGATCTTAT 758
 Db 121 CTTACCAATATCCCTGATCTTAT 180
 QY 759 TGTCTACTTTTGAACCTCCAACTGAGATGTGTCTCCACAGGCTCCACCAAGCA 818
 Db 181 TGTCTACTTTTGAACCTCCAACTGAGATGTGTCTCCACAGGCTCCACCAAGCA 240
 QY 819 GCAAGTGTATATACATCCAGAGCTGTGAGGTTTGGCAAGAGACACAGCTTCAATCCTTA 878
 Db 241 GCAAGTGTATATACATCCAGAGCTGTGAGGTTTGGCAAGAGACACAGCTTCAATCCTTA 300

QY 879 TGTAGAGGTGCCCATTTGGCTGTGAGCGAGTGGGGTGGAGTACCGCTGTGACAGGCTGC 938
DB 301 TGTAAAGGTGCCCATTTGGCTGTGAGCGAGTGGGGTGGAGTACCGCTGTGACAGGCTGC 360
QY 939 CTACCTGTTCAGAAAGGGGGGGGGCTGTGAGCGAGTGGGGTGGAGTACCGCTGTGAGTGA 998
DB 361 CTACCTGTTCAGAAAGGGGGGGGGCTGTGAGCGAGTGGGGTGGAGTACCGCTGTGAGTGA 420
QY 999 CCTGTCTTCCACCGCTTCTTCCAGAGGCGAGAGGAGAAATGAAATCCCTGATGATGC 1058
DB 421 CCTGTCTTCCACCGCTTCTTCCAGAGGCGAGAGGAGAAATGAAATCCCTGATGATGC 480
QY 1059 GGGCCTGTGAGTTCATTTGATTTGAGCAGATTAATATACCGCATTAAGAGCGGCTGCAATC 1118
DB 481 GGGCCTGTGAGTTCATTTGATTTGAGCAGATTAATATACCGCATTAAGAGCGGCTGCAATC 540
QY 1119 TTGTTACCGGGGGGAGAGGAGCGCTGAGCCTGAGCTGAGTGAAGAGCATCCCTG 1178
DB 541 TTGTTACCGGGGGGAGAGGAGCGCTGAGCCTGAGCTGAGTGAAGAGCATCCCTG 600
QY 1179 CAGCAGTGGCTCTTACCATTTGAGTGAAGTACCTTCTGTGAGCTGAGCATGAATGCTCCCT 1238
DB 601 CAGCAGTGGCTCTTACCATTTGAGTGAAGTACCTTCTGTGAGCTGAGCATGAATGCTCCCT 660
QY 1239 GGGAGTGTCCGACATGAGTGGCGTGAATCCGCTTTCAGAGGAGCAGGAGCCGATGAC 1298
DB 661 GGGAGTGTCCGACATGAGTGGCGTGAATCCGCTTTCAGAGGAGCAGGAGCCGATGAC 720
QY 1299 GTCTGTGATCGATATGTTGTACAGAACCACTCTCTGAGCTTTGTGGGACCAAAAGTGG 1358
DB 721 GTCTGTGATCGATATGTTGTACAGAACCACTCTCTGAGCTTTGTGGGACCAAAAGTGG 780
QY 1359 CAAGCTGAAGAG 1371
DB 781 CAAGCTGAAGAG 793

RESULT 9
ADA43672 ID ADA43672 standard; cDNA; 2597 BP.
XX AC ADA43672;
XX DT 20-NOV-2003 (first entry)
XX DE Human cDNA encoding secreted/transmembrane polypeptide PRO34003.
XX KM ss: gene; human; PRO: secreted protein; transmembrane protein;
XX KM endothelial cell tube formation; chondrocyte cell differentiation;
XX KM microvascular endothelial cell; tumour; lung tumour; colon tumour;
XX KM breast tumour; prostate tumour; rectal tumour; kidney tumour;
XX KM liver tumour; cytosolic; vaccine.
XX OS Homo sapiens.
XX PN US2003073190-A1.
XX PD 17-APR-2003.
XX PF 09-SEP-2002; 2002US-00238283.
XX PR 01-JUL-1998; 98US-0091358P.
XX PR 02-JUN-1999; 99WO-US012252.
XX PR 20-JUL-1999; 99US-0144758P.
XX PR 25-JUL-1999; 99US-0146222P.
XX PR 28-AUG-1999; 99US-00380137.
XX PR 30-MAR-2000; 2000WO-US008439.
XX PR 02-JUN-2000; 2000WO-US015264.
XX PR 29-AUG-2001; 2001WO-US027099.
XX PR 18-JUL-2002; 2002US-00197942.
XX PA (GETH) GENENTECH INC.
XX

PI Baker KP, Eaton DL, Filvaroff E, Goddard A, Grimaldi JC,
PI Gurney AL, Smith V, Stephan JP, Watanabe CK, Wood WI, Zhang Z,
PI Fong S;
XX WPI; 2003-585304/55.
DR P-PDB; ADA43673.
XX
XX New isolated nucleic acid encoding a PRO polypeptide, e.g. PRO20080 or
PT PRO21383, useful in molecular biology, chromosome and gene mapping, in
PT generating antisense RNA and DNA, and in gene therapy.
XX
XX Claim 2; Fig 91; 352pp; English.
XX
XX The invention relates to an isolated secreted/transmembrane (PRO)
CC polypeptide, having at least 80% sequence identity to a sequence selected
CC from any one of the 57 amino acid sequences given in specification, or to
CC a sequence encoded by a nucleic acid molecule selected from any one of
CC the nucleic acids deposited under any of the ATCC accession numbers given
CC in specification, or a sequence having at least 80% identity to PRO
CC lacking its associated signal peptide, an extracellular domain of PRO
CC with or without its associated signal peptide. Also included are vectors,
CC transformed host cells, anti-PRO antibodies, the nucleic acids encoding
CC PRO, PRO fusion proteins, inducing endothelial cell tube formation (by
CC administering PRO281, PRO1560, PRO189, PRO4499, PRO6308, PRO6000,
CC PRO10275, PRO21207, PRO20993 or PRO34274 polypeptide or its agonist) and
CC an oligonucleotide probe derived from any one of the above nucleotide
CC sequences. PRO6018 polypeptide is useful for stimulating the
CC proliferation or differentiation of chondrocyte cells. PRO1313, PRO20080
CC and PRO21383 polypeptides are useful for stimulating the proliferation of
CC human microvascular endothelial cells. PRO6071, PRO4487 and PRO6006
CC polypeptides are useful for inhibiting the proliferation of human
CC microvascular endothelial cells. PRO polypeptides are useful for
CC detecting the presence of tumour in a mammal, including tumours of lung,
CC colon, breast, prostate, rectal, kidney and liver. PRO281, PRO1560,
CC PRO189, PRO4499, PRO6308, PRO6000, PRO10275, PRO20993 and
CC PRO34274 polypeptides are useful for inducing endothelial cell tube
CC formation. PRO or the antibody are useful in the preparation of a
CC medicament for treating a condition responsive to PRO polypeptide. The
CC oligonucleotide probes are useful for isolating genomic and cDNA
CC nucleotide sequences, for measuring or detecting the expression of an
CC associated gene, and as antisense probes. PRO nucleic acid is useful as a
CC hybridisation probe, in chromosome and gene mapping, in the generation of
CC antisense RNA and DNA, and for the preparation PRO polypeptides. PRO and
CC PRO nucleic acid are useful as therapeutic agents, e.g. vaccines. The
CC present sequence encodes a PRO protein.
XX
XX Sequence 2597 BP; 688 A; 570 C; 642 G; 697 T; 0 U; 0 Other;
SQ

Query Match 13.3%; Score 793; DB 8; Length 2597;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 793; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 579 CGAGTATTTTCCACCATCTTCACGCCGGAACCTGACCAAGATCTGAGCGGATGGCAT 638
DB 1 CGAGTATTTTCCACCATCTTCACGCCGGAACCTGACCAAGATCTGAGCGGATGGCAT 60
QY 639 GTTGGCGTACGTCCTTCATATGATGATGATGATGATGATGATGATGATGATGATGATGAT 698
DB 61 GTTGGCGTACGTCCTTCATATGATGATGATGATGATGATGATGATGATGATGATGATGAT 120
QY 699 CTTACCATCATCTCTGATCTTGTATCTATCTATCTATCTATCTATCTATCTATCTATCTAT 758
DB 121 CTTACCATCATCTCTGATCTTGTATCTATCTATCTATCTATCTATCTATCTATCTATCTAT 180
QY 759 TGTCTACTTTTGAACCTTCAACCTGAGTGTGTCTCCACCAAGGCTCCACCAAGGA 818
DB 181 TGTCTACTTTTGAACCTTCAACCTGAGTGTGTGTCTCCACCAAGGCTCCACCAAGGA 240
QY 819 GCAGGTGTATATCCAAAGCTCTGAGGCTTTGCAAGAGAGACACAGCTTTCAACTCTTA 878
DB 241 GCAGGTGTATATCCAAAGCTCTGAGGCTTTGCAAGAGAGACACAGCTTTCAACTCTTA 300
QY 879 TGTAGAGGTGCCCATTTGGCTGTGAGCGAGTGGGGTGGAGTACCGCTGTGACAGGCTGC 938

Db	301	TTGTAAGATGCCCATTTGAGCTGTGAACGCAATGAGGTGAAATGATACCGGCTGTGTGACAGCTGC	360
QY	939	CTACCTGTCCAAAGCGGGGCGGTGCTTGGCAAGACCTTGGAGTCCATCCAGATGATGA	998
Db	361	CTACCTGTCCAAAGCGGGGCGCGGTGCTTGGCAAGACCTTGGAGTCCATCCAGATGATGA	420
QY	999	CCTGCTCTTACACGCTCTTCTCCAAAGGGCGAAGCGGAAAATGAAATCCTGTGATGAGTTC	1058
Db	421	CTGTGCTTTCAACGGTCTCTTCCAAAGGGCGAAGAGGAAAATGAAATCCTGTGATGAGTTC	480
QY	1059	GGCCCTGTGCATCTTCAATCTTGAACAGATTAATACCGGATTTAAGAGCGGCTGCACTC	1118
Db	481	GGCCCTGTGCATCTTCAATCTTGAACAGATTAATATACCGGATTTAAGAGCGGCTGCACTC	540
QY	1119	TTGTATACCGGGCGAAGGGCAAGCTGAGCCTGAGCTCAAGTGAAGGACATCCCTTG	1178
Db	541	TTGTATACCGGGCGAAGGGCAAGCTGAGCCTGAGCTCAAGTGAAGGACATCCCTTG	600
QY	1179	CAGCAGTGGCGCTTAAACCATTTGACGATTAATCTGTGTGAGCGTGGACATAGTCTCCCTT	1238
Db	601	CAGCAGTGGCGCTTAAACCATTTGACGATTAATCTGTGTGAGCGTGGACATAGTCTCCCTT	660
QY	1239	GGAGATGTCCGACATGTGTGCGGTGGAATTCGCCCTCTTACCGAGGACAGGACCGGATAGAC	1298
Db	661	GGAGATGTCCGACATGTGTGCGGTGGAATTCGCCCTCTTACCGAGGACAGGACCGGATAGAC	720
QY	1299	GTCTGTCAATCGCATATGTCTTACAAAGAACCATCTGTGGCTTTGTGGGACACCAAAAGTGG	1358
Db	721	GTCTGTCAATCGCATATGTCTTACAAAGAACCATCTGTGGCTTTGTGGGACACCAAAAGTGG	780
QY	1359	CAAGCTGAAGAAG 1371	
Db	781	CAAGCTGAAGAAG 793	

RESULT 10	
ADA06934	
ID	ADA06934 standard; cDNA; 2597 BP.

AC ADA06934;

DT 06-NOV-2003 (First entry)

Human PRO polynucleotide #46.

KM Human, NO: gene, ss; secreted polypeptide; transmembrane polypeptide;
KM tumour, necrosis factor- α , TNF- α ; blood; chondrocyte cell; tumour
KM adrenal; lung cancer; breast; prostate; rectum; cervix; liver;
KM microvascular endothelial cell; endothelial cell tube formation;
KM bone disorder; cartilage disorder; sports injury; proteoglycan;
KM cartilage; sports-related joint problem; articular cartilage defect;
KM osteoarthritis; rheumatoid arthritis; haemoglobin-associated disorder;
KM thalassemia; immune system cell infiltration; cancer; vulnery;
KM antineutemic; osteopathic; antineutematic; antirheumatic.

OS Homo sapiens.

PN US2003068781-A1.

PD 10-APR-2003.

PF 16-SEP-2002; 2002US-00245771.

PR 03-AUG-1999; 99US-0146843P

PR 29-AUG-2001; 2001WC-US027099

XX
XX
GENERAL INVESTMENT INC

XX	Eaton D.	Filvaroff E.	Goddard A.	Grimaldi JC
PI	Paker KD			

GURNEY AL, STEPHAN V, MACDONALD WJ NOV 74 1980

PI Fong S,

DR WPI; 2003-625486/59

 NATIONAL SCIENCE FOUNDATION

PT Novel secreted and transmembrane polypeptides, PRO polypeptides useful
PT for stimulating proliferation or differentiation of chondrocyte cells and
PT inducing endothelial cell tube formation.

PS Claim 2; Fig 91; 307pp; English.

The invention relates to isolated human PRO polypeptides (secreted and transmembrane polypeptides) and the polynucleotides encoding them. The invention also relates to an antibody which specifically binds to a PRO polypeptide, a method for stimulating the release of tumour necrosis factor- α (TNF- α) from human blood, a method for stimulating the proliferation or differentiation of chondrocyte cells and a method for detecting the presence of a tumour in a mammal (e.g. adrenal, lung, colon, breast, prostate, rectal, cervical and liver tumours). The polynucleotides are useful in molecular biology, including use as hybridisation probes, in chromosome and gene mapping, in generating antisense RNA and DNA and in gene therapy. The polynucleotides may also be used in preparing PRO polypeptides by recombinant techniques and in generating either transgenic animals or knock-out animals which are useful in the development and screening of therapeutically useful reagents. The PRO polypeptides or antibodies are used in preparing a medicament for treating a condition responsive to the polypeptides or antibodies, such as tumours, for stimulating and inhibiting proliferation of human microvascular endothelial cells, for inducing endothelial cell tube formation and for treating various bone and/or cartilage disorders such as sports injuries and arthritis. PRO polypeptides which stimulate the release of proteoglycans from cartilage are useful for treating sports-related joint problems, articular cartilage defects, osteoarthritis and rheumatoid arthritis. PRO polypeptides are also useful for treating various mammalian haemoglobin-associated disorders such as various thalassemias and conditions which may benefit from enhanced local immune system cell infiltration. This sequence represents a human PRO polynucleotide of the invention.

Sequence 2597 BP; 688 A; 570 C; 642 G; 697 T; 0 U; 0 Other;

Query Match 13.9%; Score 793; DB 8; Length 2597;

Matches 793; Conservative 0; Mismatches 0; Indels 0; Gaps 0

579 CGAGTATTTTCCACCATCTCCAGCCGAACTGACCAAGACTCTGAGGCCGATGGCAI 638

Db 1 CGAGTATTTTCCCAACCATCTCCAGCCGGAAC TGACCAAGAACCTTGAGGCGAIGCAI 80

639 GTTCGGTACGCTTCCATGATGAGTTCGTCGGCTTCGATGATTAAGATCCCTTCGGAC 858

Db 61 GTTCGCGTACGTCCTCCAGAGAGTTCGTGGCCCGAGGATTAAGATCCCTCCGACAC 120

699 CTTACCACTATCCCTGACTTGGTAATCATTAGTGCTAAGTAACTTTCCTGCTT

D9 J21 C7IACCAICAI CCIGALIIGHIHCIAHIAIICIAHIOCIHOOC.....

[illegible][illegible]

0 1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18 19 20 21 22 23 24 25 26 27 28 29 30 31 32 33 34 35 36 37 38 39 40 41 42 43 44 45 46 47 48 49 50 51 52 53 54 55 56 57 58 59 60 61 62 63 64 65 66 67 68 69 70 71 72 73 74 75 76 77 78 79 80 81 82 83 84 85 86 87 88 89 90 91 92 93 94 95 96 97 98 99

2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60
61
62
63
64
65
66
67
68
69
70
71
72
73
74
75
76
77
78
79
80
81
82
83
84
85
86
87
88
89
90
91
92
93
94
95
96
97
98
99
100
101
102
103
104
105
106
107
108
109
110
111
112
113
114
115
116
117
118
119
120
121
122
123
124
125
126
127
128
129
130
131
132
133
134
135
136
137
138
139
140
141
142
143
144
145
146
147
148
149
150
151
152
153
154
155
156
157
158
159
160
161
162
163
164
165
166
167
168
169
170
171
172
173
174
175
176
177
178
179
180
181
182
183
184
185
186
187
188
189
190
191
192
193
194
195
196
197
198
199
200
201
202
203
204
205
206
207
208
209
210
211
212
213
214
215
216
217
218
219
220
221
222
223
224
225
226
227
228
229
230
231
232
233
234
235
236
237
238
239
240
241
242
243
244
245
246
247
248
249
250
251
252
253
254
255
256
257
258
259
260
261
262
263
264
265
266
267
268
269
270
271
272
273
274
275
276
277
278
279
280
281
282
283
284
285
286
287
288
289
290
291
292
293
294
295
296
297
298
299
300
301
302
303
304
305
306
307
308
309
310
311
312
313
314
315
316
317
318
319
320
321
322
323
324
325
326
327
328
329
330
331
332
333
334
335
336
337
338
339
340
341
342
343
344
345
346
347
348
349
350
351
352
353
354
355
356
357
358
359
360
361
362
363
364
365
366
367
368
369
370
371
372
373
374
375
376
377
378
379
380
381
382
383
384
385
386
387
388
389
390
391
392
393
394
395
396
397
398
399
400
401
402
403
404
405
406
407
408
409
410
411
412
413
414
415
416
417
418
419
420
421
422
423
424
425
426
427
428
429
430
431
432
433
434
435
436
437
438
439
440
441
442
443
444
445
446
447
448
449
450
451
452
453
454
455
456
457
458
459
460
461
462
463
464
465
466
467
468
469
470
471
472
473
474
475
476
477
478
479
480
481
482
483
484
485
486
487
488
489
490
491
492
493
494
495
496
497
498
499
500
501
502
503
504
505
506
507
508
509
510
511
512
513
514
515
516
517
518
519
520
521
522
523
524
525
526
527
528
529
530
531
532
533
534
535
536
537
538
539
540
541
542
543
544
545
546
547
548
549
550
551
552
553
554
555
556
557
558
559
560
561
562
563
564
565
566
567
568
569
570
571
572
573
574
575
576
577
578
579
580
581
582
583
584
585
586
587
588
589
590
591
592
593
594
595
596
597
598
599
600
601
602
603
604
605
606
607
608
609
610
611
612
613
614
615
616
617
618
619
620
621
622
623
624
625
626
627
628
629
630
631
632
633
634
635
636
637
638
639
640
641
642
643
644
645
646
647
648
649
650
651
652
653
654
655
656
657
658
659
660
661
662
663
664
665
666
667
668
669
670
671
672
673
674
675
676
677
678
679
680
681
682
683
684
685
686
687
688
689
690
691
692
693
694
695
696
697
698
699
700
701
702
703
704
705
706
707
708
709
710
711
712
713
714
715
716
717
718
719
720
721
722
723
724
725
726
727
728
729
730
731
732
733
734
735
736
737
738
739
740
741
742
743
744
745
746
747
748
749
750
751
752
753
754
755
756
757
758
759
760
761
762
763
764
765
766
767
768
769
770
771
772
773
774
775
776
777
778
779
780
781
782
783
784
785
786
787
788
789
790
791
792
793
794
795
796
797
798
799
800
801
802
803
804
805
806
807
808
809
810
811
812
813
814
815
816
817
818
819
820
821
822
823
824
825
826
827
828
829
830
831
832
833
834
835
836
837
838
839
840
841

QY 999 CCGCTCTTCAACCGCTCTTCCAGAGCCGAGAACGGAAATATCCCTGGATGATC 1058
 DB 421 CCGCTCTTCAACCGCTCTTCCAGAGCCGAGAACGGAAATATCCCTGGATGATC 480
 QY 1059 GGGCCCTGTGATCTTCACTTCTTGAAGAGATTAATGACCCGATTAAGAGCCGCTGACATC 1118
 DB 481 GGGCCCTGTGATCTTCACTTCTTGAAGAGATTAATGACCCGATTAAGAGCCGCTGACATC 540
 QY 1119 TTGTATCCGGGGCGAGGGAGCGAGCGCTGAGCCGCTGAGCGCTGAGCGAGATCCCTG 1178
 DB 541 TTGTATCCGGGGCGAGGGAGCGAGCGCTGAGCGCTGAGCGCTGAGCGAGATCCCTG 600
 QY 1179 CAGCAGTGCCTCTTCAACATTAAGAGATTAATGACCCGATTAAGAGCCGCTGACATC 1238
 DB 601 CAGCAGTGCCTCTTCAACATTAAGAGATTAATGACCCGATTAAGAGCCGCTGACATC 660
 QY 1239 GGGAGTGTCCGAGATGAGTGGTGGAAATTCGCCCTTCAAGAGAGAGAGAGAGCCGATGAC 1298
 DB 661 GGGAGTGTCCGAGATGAGTGGTGGAAATTCGCCCTTCAAGAGAGAGAGAGAGCCGATGAC 720
 QY 1299 GTCTGTATGCGATATGATGTCAAGAAACCACTCTGCTGAGCTTTGTGGGACCAAAAGTGG 1358
 DB 721 GTCTGTATGCGATATGATGTCAAGAAACCACTCTGCTGAGCTTTGTGGGACCAAAAGTGG 780
 QY 1359 CAAAGCTGAAGAG 1371
 DB 781 CAAAGCTGAAGAG 793
 RESULT 11
 ID ADO8422 standard; cDNA; 2597 BP.
 XX ADO8422;
 AC ADO8422;
 DT 06-NOV-2003 (first entry)
 XX
 DE Novel human secreted and transmembrane protein PRO34003 cDNA.
 XX
 KW ss; gene; osteopathic; antirheumatic; antiarthritic; gene therapy;
 KW cell proliferation stimulator;
 KW chondrocyte cell differentiation stimulator;
 KW secreted and transmembrane protein; PRO; human; PRO1313; PRO20080;
 KW PRO21883; human microvascular endothelial cell proliferation; PRO6071;
 KW PRO4487; PRO6006; PRO240; PRO256; PRO698; PRO1002; PRO4316; tumour;
 KW adrenal tumour; lung tumour; colon tumour; breast tumour;
 KW prostate tumour; rectal tumour; cervical tumour; liver tumour;
 KW sports-related joint problem; articular cartilage defect; osteoarthritis;
 KW rheumatoid arthritis; tissue typing.
 XX
 OS Homo sapiens.
 XX
 FN US2003068783-A1.
 PD 10-APR-2003.
 XX
 PF 16-SEP-2002; 2002US-00245883.
 XX
 PR 09-MAY-2001; 2001US-0290589P.
 PR 29-AUG-2001; 2001MO-US027099.
 PR 18-JUL-2002; 2002US-00197942.
 XX
 PA (GENT) GENENTECH INC.
 XX
 PI Baker KP, Saton DL, Filvaroff E, Goddard A, Grimaldi JC,
 PI Gunney AL, Smith V, Stephan JP, Watanabe CK, Wood WI, Zhang Z;
 PI Fong S;
 XX
 DR WPI: 2003-625488/59.
 DR P-PSDB: ADO8423.
 XX
 PT Novel isolated PRO1313, PRO20080 or PRO21883 polypeptide useful for
 stimulating proliferation of human microvascular endothelial cells, and

PT PRO6018 polypeptide useful for stimulating proliferation of chondrocyte
 PT cells.
 XX
 PS Claim 2; Fig 91; 308bp; English.
 XX
 CC The invention describes an isolated PRO (secreted and transmembrane)
 CC polypeptide (I). PRO6018 polypeptide is useful for stimulating the
 CC proliferation or differentiation of chondrocyte cells. PRO1313, PRO20080,
 CC or PRO21883 polypeptide is useful for stimulating the proliferation of
 CC human microvascular endothelial cells. PRO6071, PRO4487, or PRO6006
 CC polypeptide is useful for inhibiting the proliferation of human
 CC microvascular endothelial cells. PRO polypeptides such as PRO240, PRO256,
 CC PRO698, PRO1002, PRO4316, etc., are useful for detecting the presence of
 CC tumour in a mammal which involves comparing the level of expression of
 CC the above mentioned polypeptides in a test sample of cells taken from the
 CC mammal, and a control sample of normal cells of the same cell type, where
 CC a higher level of expression of the PRO polypeptide in the test sample as
 CC compared to the control sample is indicative of the presence of tumour in
 CC the mammal, the tumour being adrenal tumour, lung tumour, colon tumour,
 CC breast tumour, prostate tumour, rectal tumour, cervical tumour or liver
 CC tumour. PRO6018 polypeptide is useful for treating sports-related joint
 CC problems, articular cartilage defects, osteoarthritis, or rheumatoid
 CC arthritis. PRO polypeptides are useful as molecular weight markers for
 CC protein electrophoresis. (I) is also useful for screening compounds to
 CC identify those that mimic the PRO polypeptide (agonists) or prevent the
 CC effect of the PRO polypeptide (antagonists). The polynucleotide (II)
 CC encoding (I) is useful as hybridisation probes. In chromosome and gene
 CC mapping and in the generation of anti-sense RNA and DNA. PRO nucleic acid
 CC is also useful for the preparation of PRO polypeptides. The full-
 CC length native sequence of PRO gene or its portions may be used as
 CC hybridisation probes for a cDNA library to isolate the full-length PRO
 CC cDNA or to isolate still other cDNAs. Nucleotide sequences encoding PRO
 CC can also be used to construct hybridisation probes for mapping the gene
 CC which encodes that PRO and for the genetic analysis of individuals with
 CC genetic disorders. (II) encoding (I) or its modified forms can also be
 CC used to generate either transgenic animals or knockout animals which in
 CC turn are useful in the development and screening of therapeutically
 CC useful reagents. (II) encoding PRO polypeptides are also useful in gene
 CC therapy techniques to treat conditions associated with aberrant
 CC expression or activity of (I). The PRO polypeptides and nucleic acid
 CC molecules are useful for tissue typing. This sequence encodes a novel
 CC human secreted and transmembrane PRO polypeptide.
 CC
 XX
 SQ Sequence 2597 BP; 688 A; 570 C; 642 G; 697 T; 0 U; 0 Other;
 Query Match 13.9%; Score 793; DB 8; Length 2597;
 Best Local Similarity 100.0%; Pred. No. 0;
 Matches 793; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 579 CGAGTATTTTCCACATCTCCAGCGGAGAACTGACCAAGAACTCTGAGGGATGCGAT 638
 DB 1 CGAGTATTTTCCACATCTCCAGCGGAGAACTGACCAAGAACTCTGAGGGATGCGAT 60
 QY 639 GTTCGGTACGCTTTCATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 698
 DB 61 GTTCGGTACGCTTTCATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 120
 QY 699 CTTCACATCATCCCTGATCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 758
 DB 121 CTTCACATCATCCCTGATCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 180
 QY 759 TGTCTACTTTTGGACCTTCCAACTGATGATGATGATGATGATGATGATGATGATGATGATGATG 818
 DB 181 TGTCTACTTTTGGACCTTCCAACTGATGATGATGATGATGATGATGATGATGATGATGATGATG 240
 QY 819 GAGGAGTATATCATCAAGCTGATGAGCTTTGGAAGAGAGACAGAGCTTCAACTCTTA 878
 DB 241 GAGGAGTATATCATCAAGCTGATGAGCTTTGGAAGAGAGACAGAGCTTCAACTCTTA 300
 QY 879 TGTAGAGTGGCCATTTGGCTGTGAGAGCGAGTGGGGTGAATCCGCTCTGTGACAGGCTGC 938
 DB 301 TGTAGAGTGGCCATTTGGCTGTGAGAGCGAGTGGGGTGAATCCGCTCTGTGACAGGCTGC 360

QY 933 CTACCTGTCCAAAGCGGGGCGCTGCTTGGGAGGACCCCTGGAGTCCATCCAGATGATGA 998
 Db 361 CTACCTGTCCAAAGCGGGGCGCTGCTTGGGAGGACCCCTGGAGTCCATCCAGATGATGA 420
 QY 999 CCTGCTCTTCAACCGCTCTTCTCCAAAGGCGCAAGAAAGGAAATGAAATCCCTGGATGATC 1058
 Db 421 CCTGCTCTTCAACCGCTCTTCTCCAAAGGCGCAAGAAATGAAATCCCTGGATGATC 480
 QY 1059 GGGCCCTGTGATCTTCACTTGAAGCAAGATTAATGACCCCATTAAGGAGCGGCTGAGTC 1118
 Db 481 GGGCCCTGTGATCTTCACTTGAAGCAAGATTAATGACCCCATTAAGGAGCGGCTGAGTC 540
 QY 1119 TTGTTACCGGGGCGAGGCGACGCTGACCTGCGCTGCTCAAGTGAAGGACATCCCTG 1178
 Db 541 TTGTTACCGGGGCGAGGCGACGCTGACCTGCGCTGCTCAAGTGAAGGACATCCCTG 600
 QY 1179 CAGCAGTGGCTCTTAAACATTAAGCAATTAATCTTGTGCTGCTGACATGAATGCTCCCT 1238
 Db 601 CAGCAGTGGCTCTTAAACATTAAGCAATTAATCTTGTGCTGCTGACATGAATGCTCCCT 660
 QY 1239 GGGAGTGTCCGACATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1298
 Db 661 GGGAGTGTCCGACATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 720
 QY 1299 GTCTGTCACTGCAATGCTTCAAGAAACCACTCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1358
 Db 721 GTCTGTCACTGCAATGCTTCAAGAAACCACTCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 780
 QY 1359 CAAGCTGAAGAG 1371
 Db 781 CAAGCTGAAGAG 793

RESULT 12
 ADB9715
 ID ADB9715 standard; cDNA; 2597 BP.
 XX ADB9715;
 AC 04-DEC-2003 (first entry)
 DT
 DT
 XX
 DE Human PRO polynucleotide SEQ ID 91.
 XX
 KM Human; gene; se; PRO; secreted polypeptide; transmembrane polypeptide;
 KM tumour necrosis factor-alpha; TNF-alpha; blood; chondrocyte cell; tumour;
 KM adrenal; lung; colon; breast; prostate; rectum; cervix; liver; cancer;
 KM microvascular endothelial cell; endothelial cell tube formation;
 KM sports-related joint problem; articular cartilage defect; osteoarthritis;
 KM rheumatoid arthritis; osteopathic; antiinflammatory; antiarthritic.
 XX
 OS Homo sapiens.
 XX
 PN US2003082728-A1.
 XX
 PD 01-MAY-2003.
 XX
 PF 16-SEP-2002; 2002US-00245185.
 XX
 PR 07-JUL-1998; 98US-00919782.
 PR 02-JUN-1999; 99WO-0501252.
 PR 25-AUG-1999; 99US-00380137.
 PR 29-AUG-2001; 2001WO-05027099.
 PR 18-JUL-2002; 2002US-00197942.
 XX
 PA (GENTH) GENENTECH INC.
 XX
 PI Baker KP, Baton DL, Filvaroff E, Goddard A, Grimaldi JC,
 PI Gurey AL, Smith V, Stephan JP, Watanabe CK, Wood WT, Zhang Z,
 PI Fong S,
 XX WPI: 2003-743698/70.
 DR P-PsDB, ADB9716.
 XX

PT New isolated, secreted and transmembrane PRO polypeptides and nucleic
 PT acids, useful for diagnosing, preventing and/or treating tumors, such as
 PT lung, colon, breast, prostate, rectal, kidney or liver tumors.
 XX
 PS Claim 2; Fig 91; 308pp; English.
 XX
 CC The invention relates to isolated human PRO polypeptides (secreted and
 CC transmembrane polypeptides) and the polynucleotides encoding them. The
 CC invention also relates to an antibody which specifically binds to a PRO
 CC polypeptide, a method for stimulating the release of tumour necrosis
 CC factor-alpha (TNF-alpha) from human blood, a method for stimulating the
 CC proliferation or differentiation of chondrocyte cells and a method for
 CC detecting the presence of a tumour in a mammal (e.g. adrenal, lung,
 CC colon, breast, prostate, rectal, cervical and liver tumors). The
 CC polynucleotides are useful in molecular biology, including uses as
 CC hybridisation probes, in chromosome and gene mapping, in generating
 CC antisense RNA and DNA and in gene therapy. The polynucleotides may also
 CC be used in preparing PRO polypeptides by recombinant techniques and in
 CC generating either transgenic animals or knock-out animals which are
 CC useful in the development and screening of therapeutically useful
 CC reagents. The PRO polypeptides or antibodies are used in preparing a
 CC medicament for treating a condition responsive to the polypeptides or
 CC antibodies, such as tumours, for stimulating and inhibiting proliferation
 CC of human microvascular endothelial cells, for inducing endothelial cell
 CC tube formation and for treating sports-related joint problems, articular
 CC cartilage defects, osteoarthritis and rheumatoid arthritis. This sequence
 CC represents a human PRO polynucleotide of the invention.
 CC
 SQ Sequence 2597 BP; 688 A; 570 C; 642 G; 697 T; 0 U; 0 Other;
 Query Match 13.9%; Score 793; DB 8; Length 2597;
 Best Local Similarity 100.0%; Pred. No. 0;
 Matches 793; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 579 CGAGTATTTTCCACCATCTCCAGCCGGAACCTGACCAAGACTCTGAGGGGATGGAT 638
 Db 1 CGAGTATTTTCCACCATCTCCAGCCGGAACCTGACCAAGACTCTGAGGGGATGGAT 60
 QY 639 GTTGGGATGACCTTCCATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 698
 Db 61 GTTGGGATGACCTTCCATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 120
 QY 699 CTTCACCATCATCCTGATCTTATGATGATGATGATGATGATGATGATGATGATGATGATGAT 758
 Db 121 CTTCACCATCATCCTGATCTTATGATGATGATGATGATGATGATGATGATGATGATGATGAT 180
 QY 759 TGTCTACTTTTGAACCTCCCAACCGATGATGATGATGATGATGATGATGATGATGATGATGAT 818
 Db 181 TGTCTACTTTTGAACCTCCCAACCGATGATGATGATGATGATGATGATGATGATGATGATGAT 240
 QY 819 GCAAGTGTATATACATCAAGCTGTGAGGCTTTGCAAGAGACACAGCCTTCACTCTTA 878
 Db 241 GCAAGTGTATATACATCAAGCTGTGAGGCTTTGCAAGAGACACAGCCTTCACTCTTA 300
 QY 879 TGTAGAGTGGCCATTTGCTGAGAGCGAGTGGGATGATGATGATGATGATGATGATGATGAT 938
 Db 301 TGTAGAGTGGCCATTTGCTGAGAGCGAGTGGGATGATGATGATGATGATGATGATGATGAT 360
 QY 939 CTACCTGTCCAAAGCGGGGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 998
 Db 361 CTACCTGTCCAAAGCGGGGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 420
 QY 999 CCTGCTCTTCAACCGCTCTTCTCCAAAGGCGCAAGAAATGAAATCCCTGGATGATC 1058
 Db 421 CCTGCTCTTCAACCGCTCTTCTCCAAAGGCGCAAGAAATGAAATCCCTGGATGATC 480
 QY 1059 GGGCCCTGTGATCTTCACTTGAAGCAAGATTAATGACCCCATTAAGGAGCGGCTGAGTC 1118
 Db 481 GGGCCCTGTGATCTTCACTTGAAGCAAGATTAATGACCCCATTAAGGAGCGGCTGAGTC 540
 QY 1119 TTGTTACCGGGGCGAGGCGACGCTGACCTGCGCTGCTCAAGTGAAGGACATCCCTG 1178
 Db 541 TTGTTACCGGGGCGAGGCGACGCTGACCTGCGCTGCTCAAGTGAAGGACATCCCTG 600

QY 1179 CAGAGTGGCGCTCTTAAACCATGAGATACCTTGTGGCTGACATGATGCTCCCT 1238
 Db 601 CAGAGTGGCGCTCTTAAACCATGAGATACCTTGTGGCTGACATGATGCTCCCT 660
 QY 1239 GGGAGTGGCGCATGATGCTGAGATGCTGCGCTTACAGAGAGACAGGACCGCATGAC 1298
 Db 661 GGGAGTGGCGCATGATGCTGAGATGCTGCGCTTACAGAGAGACAGGACCGCATGAC 720
 QY 1299 GTCTGTCATGCAATGCTCTACAGAACCACTCTCTGCTTGTGGGACCAAAAGTGG 1358
 Db 721 GTCTGTCATGCAATGCTCTACAGAACCACTCTCTGCTTGTGGGACCAAAAGTGG 780
 QY 1359 CAGCTGAAGAAG 1371
 Db 781 CAGCTGAAGAAG 793
 RESULT 13
 ADB86998
 ID ADB86998 standard; cDNA; 2597 BP.
 AC ADB86998;
 XX 04-DEC-2003 (first entry)
 DT Human PRO polynucleotide #46.
 XX
 KM Human, gene; *se*, PRO; secreted polypeptide; transmembrane polypeptide;
 KM tumour necrosis factor- α ; TNF- α ; blood; chondrocyte cell; tumour;
 KM adrenal; lung; colon; breast; prostate; rectum; cervix; liver; cancer;
 KM microvascular endothelial cell; endothelial cell tube formation;
 KM sports-related joint problem; articular cartilage defect; osteoarthritis;
 KM rheumatoid arthritis; osteopathic; antirheumatic; antiarthritic.
 XX
 OS Homo sapiens.
 XX
 PN US2003082726-A1.
 XX 01-MAY-2003.
 PD 12-SEP-2002; 2002US-00243095.
 PF 29-AUG-2001; 2001NO-US027099.
 PR 18-JUL-2002; 2002US-00197942.
 XX
 PA (GENENTECH INC.
 XX Baker KP, Batton DL, Filvaroff E, Goddard A, Grimaldi JC;
 PI Gurney AL, Smith V, Stephan JP, Watanabe CK, Wood WI, Zhang Z;
 PI Fong S;
 XX WPI; 2003-743897/70.
 DR P-PSDB; ADB86999.
 XX
 PT New isolated, secreted and transmembrane PRO polypeptides and nucleic
 PT acids, useful for diagnosing, preventing and/or treating tumors, such as
 PT lung, colon, breast, prostate, rectal, kidney or liver tumors.
 XX
 PS Claim 2; Fig 91; 308pp; English.
 CC The invention relates to isolated human PRO polypeptides (secreted and
 CC transmembrane polypeptides) and the polynucleotides encoding them. The
 CC invention also relates to an antibody which specifically binds to a PRO
 CC polypeptide, a method for stimulating the release of tumour necrosis
 CC factor- α (TNF- α) from human blood, a method for stimulating the
 CC proliferation or differentiation of chondrocyte cells and a method for
 CC detecting the presence of a tumour in a mammal (e.g. adrenal, lung,
 CC colon, breast, prostate, rectal, cervical and liver tumours). The
 CC polynucleotides are useful in molecular biology, including uses as
 CC hybridisation probes, in chromosome and gene mapping, in generating
 CC antisense RNA and DNA in gene therapy. The polynucleotides may also
 CC be used in preparing PRO polypeptides by recombinant techniques and in

CC generating either transgenic animals or knock-out animals which are
 CC useful in the development and screening of therapeutically useful
 CC reagents. The PRO polypeptides or antibodies are used in preparing a
 CC medicament for treating a condition responsive to the polypeptides or
 CC antibodies, such as tumours, for stimulating and inhibiting proliferation
 CC of human microvascular endothelial cells, for inducing endothelial cell
 CC tube formation and for treating sports-related joint problems, articular
 CC cartilage defects, osteoarthritis and rheumatoid arthritis. This sequence
 CC represents a human PRO polynucleotide of the invention.
 XX
 SQ Sequence 2597 BP; 688 A; 570 C; 642 G; 697 T; 0 U; 0 Other;
 Query Match 13.9%; Score 793; DB 8; Length 2597;
 Best Local Similarity 100.0%; Pred. No. 0;
 Matches 793; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 579 CGAGTATTTTCCACCATCTCCAGCCGGAACATGACCAAGAACTCGAGGCGGATGGCAT 638
 Db 1 CGAGTATTTTCCACCATCTCCAGCCGGAACATGACCAAGAACTCGAGGCGGATGGCAT 60
 QY 639 GTTCGCGTACGCTCTTCCATGATGATGCTGCTGCTGATGATTAAGATCCCTTCGACAC 698
 Db 61 GTTCGCGTACGCTCTTCCATGATGATGCTGCTGCTGATGATTAAGATCCCTTCGACAC 120
 QY 699 CTTCACCATCATCCCTGACCTTTGATATCTATCTATGCTTATAGAGAGTGGCAACTT 758
 Db 121 CTTCACCATCATCCCTGACCTTTGATATCTATCTATGCTTATAGAGAGTGGCAACTT 180
 QY 759 TGTCTACTTTTGAACCTCCACCTGAGATGATGATGATGATGATGATGATGATGATGATG 818
 Db 181 TGTCTACTTTTGAACCTCCACCTGAGATGATGATGATGATGATGATGATGATGATGATG 240
 QY 819 GCAGGTGATATCATCCAGCTCTGAGAGCTTTGCAAGAGAGACAGCCCTCACTCCTTA 878
 Db 241 GCAGGTGATATCATCCAGCTCTGAGAGCTTTGCAAGAGAGACAGCCCTCACTCCTTA 300
 QY 879 TGTAGAGTGGCCCATTTGGCTGTGAGCGCAGTGGGATGAGTACCGCTCTGAGAGCTGC 938
 Db 301 TGTAGAGTGGCCCATTTGGCTGTGAGCGCAGTGGGATGAGTACCGCTCTGAGAGCTGC 360
 QY 939 CTACCTGTCCAAAGCGGCGGCGCTGCTGCGAGAGACCTTTGAGTGCATCAATGATGA 998
 Db 361 CTACCTGTCCAAAGCGGCGGCGCTGCTGCGAGAGACCTTTGAGTGCATCAATGATGA 420
 QY 999 CCTGCTCTTACACGCTCTTCTCCAAAGGCGCAAGCGGAAATGAAATCCCTGATGATGC 1058
 Db 421 CCTGCTCTTACACGCTCTTCTCCAAAGGCGCAAGCGGAAATGAAATCCCTGATGATGC 480
 QY 1059 GGCCCTGTGATCTTCACTTTGAAGCAGATTAATGACCGCATTAAGAGGCGCTGCAGTC 1118
 Db 481 GGCCCTGTGATCTTCACTTTGAAGCAGATTAATGACCGCATTAAGAGGCGCTGCAGTC 540
 QY 1119 TTGTTACCGGGGAGAGGAGCGCTGAGACCTGGCTGCTCAAGGTGAAGACATCCCTCTG 1178
 Db 541 TTGTTACCGGGGAGAGGAGCGCTGAGACCTGGCTGCTCAAGGTGAAGACATCCCTCTG 600
 QY 1179 CAGAGTGGCGCTCTTAAACCATGAGATACCTTGTGGCTGACATGATGCTCCCT 1238
 Db 601 CAGAGTGGCGCTCTTAAACCATGAGATACCTTGTGGCTGACATGATGCTCCCT 660
 QY 1239 GGGAGTGGCGCATGATGCTGAGATGCTGCGCTTACAGAGAGACAGGACCGCATGAC 1298
 Db 661 GGGAGTGGCGCATGATGCTGAGATGCTGCGCTTACAGAGAGACAGGACCGCATGAC 720
 QY 1299 GTCTGTCATGCAATGCTCTACAGAACCACTCTCTGCTTGTGGGACCAAAAGTGG 1358
 Db 721 GTCTGTCATGCAATGCTCTACAGAACCACTCTCTGCTTGTGGGACCAAAAGTGG 780
 QY 1359 CAGCTGAAGAAG 1371
 Db 781 CAGCTGAAGAAG 793

RESULT 14
 ADB66153
 ID ADB66153 standard; cDNA; 2597 BP.
 XX ADB66153;
 AC ADB66153;
 DT 04-DEC-2003 (first entry)
 XX
 DE Human cDNA encoding secreted/transmembrane polypeptide PRO34003.
 XX
 KW bs; gene; human; PRO; secreted protein; transmembrane protein;
 KW endothelial cell tube formation; chondrocyte cell differentiation;
 KW microvascular endothelial cell; tumour; lung tumour; colon tumour;
 KW breast tumour; prostate tumour; rectal tumour; kidney tumour;
 KW liver tumour; cytostatic; vaccine.
 XX
 OS Homo sapiens.
 XX
 XX US2003082729-A1.
 XX
 XX 01-MAY-2003.
 XX
 XX 16-SEP-2002; 2002US-00245427.
 XX
 XX 02-FEB-2001; 2001US-0266421P.
 XX
 XX 09-FEB-2001; 2001US-0267623P.
 XX
 XX 29-AUG-2001; 2001WO-US027099.
 XX
 XX 18-JUL-2002; 2002US-00197942.
 XX
 XX (GERTH) GENENTECH INC.
 XX
 XX Baker KP, Eaton DL, Filvaroff E, Goddard A, Grimaldi JC,
 XX Gurney AL, Smith V, Stephan JP, Watanabe CK, Wood WL, Zhang Z,
 XX Fong S,
 XX
 XX WPI: 2003-786917/74.
 XX
 XX P-PSDB: ADB66154.
 XX
 XX New isolated, secreted and transmembrane PRO polypeptides and nucleic
 XX acids, useful for diagnosing, preventing and/or treating tumors, such as
 XX lung, colon, breast, prostate, rectal, kidney or liver tumors.
 XX
 XX Claim 2; Fig 91; 308pp; English.
 XX
 XX The invention relates to an isolated secreted/transmembrane (PRO)
 XX polypeptide, having at least 80% sequence identity to a sequence selected
 XX from any one of the 57 amino acid sequences given in specification, or to
 XX a sequence encoded by a nucleic acid molecule selected from any one of
 XX the nucleic acids deposited under any of the ATCC accession numbers given
 XX in specification, or a sequence having at least 80% identity to PRO
 XX lacking its associated signal peptide, an extracellular domain of PRO
 XX with or without its associated signal peptide. Also included are vectors,
 XX transformed host cells, anti-PRO antibodies, the nucleic acids encoding
 XX PRO, PRO fusion proteins, inducing endothelial cell tube formation (by
 XX administering PRO281, PRO1560, PRO189, PRO4499, PRO6000,
 XX PRO10275, PRO21037, PRO20933 or PRO34274 polypeptide or its agonist) and
 XX an oligonucleotide probe derived from any one of the above nucleotide
 XX sequences. PRO6018 polypeptide is useful for stimulating the
 XX proliferation or differentiation of chondrocyte cells. PRO1313, PRO20080
 XX and PRO21383 polypeptides are useful for stimulating the proliferation of
 XX human microvascular endothelial cells. PRO6071, PRO4487 and PRO6006
 XX polypeptides are useful for inhibiting the proliferation of human
 XX microvascular endothelial cells. PRO polypeptides are useful for
 XX detecting the presence of tumour in a mammal, including tumours of lung,
 XX colon, breast, prostate, rectal, kidney and liver. PRO281, PRO1560,
 XX PRO189, PRO4499, PRO6000, PRO10275, PRO21037, PRO20933 and
 XX PRO34274 polypeptides are useful for inducing endothelial cell tube
 XX formation. PRO or the antibody are useful in the preparation of a
 XX medicament for treating a condition responsive to PRO polypeptide. The
 XX oligonucleotide probes are useful for isolating genomic and cDNA
 XX nucleotide sequences, for measuring or detecting the expression of an
 XX associated gene, and as antisense probes. PRO nucleic acid is useful as a
 XX hybridisation probe, in chromosome and gene mapping, in the generation of

CC antisense RNA and DNA, and for the preparation PRO polypeptides. PRO and
 CC PRO nucleic acid are useful as therapeutic agents, e.g. vaccines. The
 CC present sequence encodes a PRO protein.
 XX
 XX Sequence 2597 BP; 688 A; 570 C; 642 G; 697 T; 0 U; 0 Other;
 SQ
 Query Match 13.9%; Score 793; DB 8; Length 2597;
 Best Local Similarity 100.0%; Pred. No. 0;
 Matches 793; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 579 CGAGTATTTTCCACATCTCCAGCCGGAACCTGACCAAGACTCTGAGGGATGGCAT 638
 DB 1 CGAGTATTTTCCACATCTCCAGCCGGAACCTGACCAAGACTCTGAGGGATGGCAT 60
 QY 639 GTTGCAGTACGCTCTCCATGATGAGTTGTGGCTCGATGATTAAGATCCCTTGCAAC 698
 DB 61 GTTGCAGTACGCTCTCCATGATGAGTTGTGGCTCGATGATTAAGATCCCTTGCAAC 120
 QY 699 CTTACACATCATCCCTGACCTTATATCTATATCTATGCTTATGCTTTAGAGTGGCACTT 758
 DB 121 CTTACACATCATCCCTGACCTTATATCTATATCTATGCTTATGCTTTAGAGTGGCACTT 180
 QY 759 TGTCTACTTTTGAACCTCCCAACCTGAGATGATGCTGCTCCACAGGCTCCACCAAGGA 818
 DB 181 TGTCTACTTTTGAACCTCCCAACCTGAGATGATGCTGCTCCACAGGCTCCACCAAGGA 240
 QY 819 GCAGGTGATACATCAAGCTGTGAGGCTTTGCAAGAGGACACAGCTTCAACTCCTA 878
 DB 241 GCAGGTGATACATCAAGCTGTGAGGCTTTGCAAGAGGACACAGCTTCAACTCCTA 300
 QY 879 TGTAGAGGTGCCCATTTGGCTGTGAGCGGAGTGGAGTACCGGCTGTGAGGCTGC 938
 DB 301 TGTAGAGGTGCCCATTTGGCTGTGAGCGGAGTGGAGTACCGGCTGTGAGGCTGC 360
 QY 939 CTACCTGTCCAAAGCGGGGCGGTGCTTGGCAGGACCTTTGAGTCCATCCAGATGATGA 998
 DB 361 CTACCTGTCCAAAGCGGGGCGGTGCTTGGCAGGACCTTTGAGTCCATCCAGATGATGA 420
 QY 999 CCTGCTCTTCCACCGCTCTTCCAGAGGCGCAAGCGGAAATGAAATCCCTGAGATGATC 1058
 DB 421 CCTGCTCTTCCACCGCTCTTCCAGAGGCGCAAGCGGAAATGAAATCCCTGAGATGATC 480
 QY 1059 GGCCCTGTGATCTTCACTTGAAGCAGATTAATGACCGCATTAAGAGCGGCTGACGTC 1118
 DB 481 GGCCCTGTGATCTTCACTTGAAGCAGATTAATGACCGCATTAAGAGCGGCTGACGTC 540
 QY 1119 TTGTTACCGGGCGGAGGCGCAGCTGAGACTTGCTGCTCAAGTGAAGGACATCCCTG 1178
 DB 541 TTGTTACCGGGCGGAGGCGCAGCTGAGACTTGCTGCTCAAGTGAAGGACATCCCTG 600
 QY 1179 CAGCAGTGGCTCTTAACCATTTGACGATTAATCTGTTGGCTGAGCATGAATGCTCCCT 1238
 DB 601 CAGCAGTGGCTCTTAACCATTTGACGATTAATCTGTTGGCTGAGCATGAATGCTCCCT 660
 QY 1239 GGAAGTGTCCGACATGATGCTGCTGATTCCTGCTTTCACGAGAGACAGGACCGCATGAC 1298
 DB 661 GGAAGTGTCCGACATGATGCTGCTGATTCCTGCTTTCACGAGAGACAGGACCGCATGAC 720
 QY 1299 GTCTGTACGATGATGCTGCTGATTCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1358
 DB 721 GTCTGTACGATGATGCTGCTGATTCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 780
 QY 1359 CAACTGAAGAG 1371
 DB 781 CAACTGAAGAG 793
 RESULT 15
 ADB99831
 ID ADB99831 standard; cDNA; 2597 BP.
 XX
 XX ADB99831;
 AC ADB99831;
 XX

XX	04-DEC-2003 (first entry)
DE	Human PRO polynucleotide SEQ ID 91.
KW	Human; gene; ss; PRO; secreted polypeptide; transmembrane polypeptide;
KM	tumour necrosis factor-alpha; TNF-alpha; blood; chondrocyte cell; tumour;
KM	adrenal; lung; colon; breast; prostate; rectum; cervix; liver; cancer;
KM	microvascular endothelial cell; endothelial cell tube formation;
KM	sports-related joint problem; articular cartilage defect; osteoarthritis;
KM	rheumatoid arthritis; osteopathic; anti-rheumatic; antiarthritic.
OS	Homo sapiens.
PN	US2003073192-A1.
PD	17-APR-2003.
PF	16-SEP-2002; 2002US-00245055.
PR	20-JUL-1999; 99US-0144758P.
PR	28-JUL-1999; 99US-0146222P.
PR	30-MAR-2000; 2000WO-US008439.
PR	02-JUN-2000; 2000WO-US015264.
PR	29-AUG-2001; 2001WO-US027099.
PR	18-JUL-2002; 2002US-00197942.
PA	(GENTH) GENENTECH INC.
PI	Baker KP, Eaton DL, Filvaroff E, Goddard A, Grimaldi JC,
PI	Gurney AL, Smith V, Stephan JP, Watanabe CK, Wood WT, Zhang Z;
PI	Fong S;
WP	WP1; 2003-743813/70.
P-PDB	P-PDB; ADB99832.
PT	New isolated secreted and transmembrane PRO polypeptides, e.g. PRO20080
PT	or PRO21383, useful in stimulating the proliferation or differentiation
PT	of chondrocyte cells and detecting the presence of a tumor in a mammal.
PS	Claim 2, Fig 91; 30pp; English.
CC	The invention relates to isolated human PRO polypeptides (secreted and
CC	transmembrane polypeptides) and the polynucleotides encoding them. The
CC	invention also relates to an antibody which specifically binds to a PRO
CC	polypeptide, a method for stimulating the release of tumour necrosis
CC	factor-alpha (TNF-alpha) from human blood, a method for stimulating the
CC	proliferation or differentiation of chondrocyte cells and a method for
CC	detecting the presence of a tumour in a mammal (e.g. adrenal, lung,
CC	colon, breast, prostate, rectal, cervical and liver tumours). The
CC	polynucleotides are useful in molecular biology, including uses as
CC	hybridisation probes, in chromosome and gene mapping, in generating
CC	antisense RNA and DNA and in gene therapy. The polynucleotides may also
CC	be used in preparing PRO polypeptides by recombinant techniques and in
CC	generating either transgenic animals or knock-out animals which are
CC	useful in the development and screening of therapeutically useful
CC	reagents. The PRO polypeptides or antibodies are used in preparing a
CC	medicament for treating a condition responsive to the polypeptides or
CC	antibodies, such as tumours, for stimulating and inhibiting proliferation
CC	of human microvascular endothelial cells, for inducing endothelial cell
CC	tube formation and for treating sports-related joint problems, articular
CC	cartilage defects, osteoarthritis and rheumatoid arthritis. This sequence
CC	represents a human PRO polynucleotide of the invention.
SQ	Sequence 2597 BP; 688 A; 570 C; 642 G; 697 T; 0 U; 0 Other;
DY	Query Match 13.9%; Score 793; DB 9; Length 2597; Best Local Similarity 100.0%; Pred. No. 0; Matches 793; Conservative 0; Mismatches 0; Indels 0; Gaps 0
DB	579 CGAGTATTTTCACCATCTCCAGCGCGAAATGACGAAGAATCCTGAGCGGATGGCAT 638 1 CGAGTATTTTCACCATCTCCAGCGCGAAATGACGAAGAATCCTGAGCGGATGGCAT 60

QY	639	GTTCGGCTAGCTCTTCCATATAGTATCGCGGCTGCATATTAAGATCCCTTGCGACAC	639
Db	61	GTTCGGCTAGCTCTTCCATATAGTATCGCGGCTGCATATTAAGATCCCTTGCGACAC	120
QY	699	CTTCCACCATCATCCCTGACTTTTGATATATCTATGTCTATAGTCTTTAGCATGCGCACTT	758
Db	121	CTTCCACCATCATCCCTGACTTTTGATATATCTATGTCTATAGTCTTTAGCATGCGCACTT	180
QY	759	TGTCTATCTTTTGAACCTCCCAACCTGAGATGATGTCTCCACAGAGCTCCACACCAAGAA	818
Db	181	TGTCTATCTTTTGAACCTCCCAACCTGAGATGATGTCTCTCCACAGAGCTCCACACCAAGAA	240
QY	819	GCAAGTGTATATCATCCAAAGCTCGTAGAGCTTTGCAAGAGAGACACAGCTTTCAACTCTTA	878
Db	241	GCAAGTGTATATCATCCAAAGCTCGTAGAGCTTTGCAAGAGAGACACAGCTTTCAACTCTTA	300
QY	879	TGTAGAGATGCGCCATTGAGCTGTGAGCGCAGTGGGGGTGAGATACCGCTGTGTCAGAGCTGC	938
Db	301	TGTAGAGATGCGCCATTGAGCTGTGAGCGCAGTGGGGGTGAGATACCGCTGTGTCAGAGCTGC	360
QY	939	CTACCTGTCCAAAGCGGGGCGGTGCTTGGCAGAGACCTTTGAGTCCATCCATATATGA	998
Db	361	CTACCTGTCCAAAGCGGGGCGGTGCTTGGCAGAGACCTTTGAGTCCATCCATATATGA	420
QY	999	CCGCTCTTCAACCGTCTTCTCCAAAGGCGCAAGACGGAAATGAATCCCTGATGAGTGC	1058
Db	421	CCGCTCTTCAACCGTCTTCTCCAAAGGCGCAAGACGGAAATGAATCCCTGATGAGTGC	480
QY	1059	GGCCCTGTGCATCTTTCATCTTGAAGACATTAATATGACCGCATTAAGAGCGGGCTGCAGTC	1118
Db	481	GGCCCTGTGCATCTTTCATCTTGAAGACATTAATATGACCGCATTAAGAGCGGGCTGCAGTC	540
QY	1119	TTGTATCCGGGGCGAGAGGACGCTGGACCTGGGCTCAAGGTGAAGGACATCCCTGT	1178
Db	541	TTGTATCCGGGGCGAGAGGACGCTGGACCTGGGCTCAAGGTGAAGGACATCCCTGT	600
QY	1179	CAGCAGTGCCTCTTAAACATTAAGCATAACTCTGTGAGCTGGACATGAATGCTCCCT	1238
Db	601	CAGCAGTGCCTCTTAAACATTAAGCATAACTCTGTGAGCTGGACATGAATGCTCCCT	660
QY	1239	GGAGATGTCCGAATAGTGGGTGAATTCGCCGTCTTACGAGAGGACAGAGGCCCATGAC	1298
Db	661	GGAGATGTCCGAATAGTGGGTGAATTCGCCGTCTTACGAGAGGACAGAGGCCCATGAC	720
QY	1299	GTCTGTATGCGCATATGTCTTACAAAGAACCACTCTGTGGCTTTGTGGGACCAAAAAGTGG	1358
Db	721	GTCTGTATGCGCATATGTCTTACAAAGAACCACTCTGTGGCTTTGTGGGACCAAAAAGTGG	780
QY	1359	CAAGCTGAAGAAG 1371	
Db	781	CAAGCTGAAGAAG 793	

Search completed: February 20, 2004, 03:12:05
Job time : 1381 secs

QY 579 CGAGTATTTTCCACCATCTCCAGCCGAAACTGACCAAGAACTCTGAGGCGGATGGCAT 638
 |||||
 DB 1 CGAGTATTTTCCACCATCTCCAGCCGAAACTGACCAAGAACTCTGAGGCGGATGGCAT 60

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: February 20, 2004, 02:26:30 : Search time 260 Seconds

(without alignments)
12147.021 Million cell updates/sec

Title: US-09-964-956-12
Perfect score: 5691
Sequence: 1 atgaagacatgcctcgaa.....gctagacatggaataaa 5691

Scoring table: OLIGO_NUC
Gapop 60.0 , Gapext 60.0

Searched: 682709 seqs, 277475446 residues

Word size : 0

Total number of hits satisfying chosen parameters: 1365418

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Listing first 1000 summaries

Database :

Issued Parents NA: *
1: /cgn2_6/prodata/2/ina/5A_COMB.seq:*
2: /cgn2_6/prodata/2/ina/5B_COMB.seq:*
3: /cgn2_6/prodata/2/ina/6A_COMB.seq:*
4: /cgn2_6/prodata/2/ina/6B_COMB.seq:*
5: /cgn2_6/prodata/2/ina/PCFUS_COMB.seq:*
6: /cgn2_6/prodata/2/ina/Backfiles1.seq:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	* Query Match Length	DB ID	Description
C 1	20	0.4	249 4	US-09-252-991A-3017 Sequence 3017, Ap
C 2	20	0.4	311 4	US-09-313-294A-7021 Sequence 7021, Ap
C 3	20	0.4	462 4	US-09-401-064-165 Sequence 165, App
C 4	20	0.4	616 3	US-09-385-982-238 Sequence 238, App
C 5	20	0.4	894 4	US-09-252-991A-2754 Sequence 2754, Ap
C 6	20	0.4	1029 4	US-09-252-991A-2651 Sequence 2651, Ap
C 7	20	0.4	1455 4	US-09-252-991A-3212 Sequence 3212, Ap
C 8	20	0.4	1456 4	US-09-576-594-779 Sequence 779, App
C 9	20	0.4	1636 4	US-08-530-915A-31 Sequence 31, App
C 10	20	0.4	2522 4	US-09-620-312D-702 Sequence 702, App
C 11	20	0.4	13842 3	US-09-105-537-5 Sequence 5, App
C 12	20	0.4	36778 3	US-09-320-878-19 Sequence 19, App
C 13	20	0.4	38506 3	US-09-141-908-11 Sequence 11, App
C 14	20	0.4	38506 4	US-09-457-440-19 Sequence 19, App
C 15	20	0.4	38506 4	US-09-657-440-19 Sequence 19, App
C 16	19	0.3	531 5	PCT-US95-04583-1 Sequence 1, App
C 17	19	0.3	588 4	US-08-924-695A-3 Sequence 3, App
C 18	19	0.3	791 4	US-09-657-346A-10 Sequence 10, App
C 19	19	0.3	900 1	US-07-955-041-1 Sequence 1, App
C 20	19	0.3	900 1	US-08-227-455-1 Sequence 1, App
C 21	19	0.3	900 1	US-08-472-482-1 Sequence 1, App
C 22	19	0.3	900 1	US-08-487-069-1 Sequence 1, App
C 23	19	0.3	1236 4	US-09-328-952-3326 Sequence 3326, Ap
C 24	19	0.3	1407 4	US-09-894-844-35 Sequence 35, App
C 25	19	0.3	2238 1	US-07-841-651-1 Sequence 1, App
C 26	19	0.3	2477 4	US-09-907-794A-169 Sequence 169, App
C 27	19	0.3	2477 4	US-09-905-125A-169 Sequence 169, App

101	-17	0.3	401	4	US-09-339-338-196	Sequence 196, App	C 174	17	0.3	148	4	US-09-485-632B-1	Sequence 1, Appl
102	-17	0.3	401	4	US-09-433-826B-196	Sequence 196, App	C 175	17	0.3	153	4	US-09-472-112-2	Sequence 2, Appl
103	-17	0.3	401	4	US-09-604-287A-196	Sequence 196, App	C 176	17	0.3	160	4	US-09-485-632B-7	Sequence 7, Appl
104	-17	0.3	401	4	US-09-834-759-196	Sequence 196, App	C 177	17	0.3	177	4	US-09-252-991A-11683	Sequence 11683, A
105	-17	0.3	402	4	US-09-543-681A-3530	Sequence 3530, Ap	C 178	17	0.3	1469	1	US-08-053-867A-3	Sequence 3, Appl
106	-17	0.3	417	4	US-09-621-976-1344	Sequence 1344, Ap	C 179	17	0.3	1469	1	US-08-053-867A-6	Sequence 6, Appl
107	-17	0.3	423	4	US-09-252-991A-13079	Sequence 13079, A	C 180	17	0.3	1512	4	US-09-540-235-1742	Sequence 1742, Ap
108	-17	0.3	453	4	US-09-621-976-1518	Sequence 1518, Ap	C 181	17	0.3	1337	2	US-09-252-991A-11555	Sequence 11555, A
109	-17	0.3	474	4	US-09-621-976-17384	Sequence 17384, A	C 182	17	0.3	1237	2	US-08-839-006-1	Sequence 1, Appl
110	-17	0.3	481	4	US-09-621-976-17383	Sequence 17383, A	C 183	17	0.3	1339	2	US-09-089-588-55	Sequence 55, Appl
111	-17	0.3	484	4	US-09-621-976-2660	Sequence 2660, Ap	C 184	17	0.3	1569	2	US-08-901-306-3	Sequence 3, Appl
112	-17	0.3	486	4	US-09-621-976-736	Sequence 736, App	C 185	17	0.3	1569	2	US-09-180-271-3	Sequence 3, Appl
113	-17	0.3	497	4	US-09-833-381-1307	Sequence 1307, Ap	C 186	17	0.3	1575	4	US-09-252-991A-1521	Sequence 1521, Ap
114	-17	0.3	500	4	US-08-808-303-11	Sequence 11, Appl	C 187	17	0.3	1643	4	US-09-599-360B-43	Sequence 43, Appl
115	-17	0.3	505	3	US-08-996-533-11	Sequence 15835, A	C 188	17	0.3	1643	4	US-09-023-655-753	Sequence 753, App
116	-17	0.3	510	4	US-09-252-991A-15535	Sequence 3443, Ap	C 189	17	0.3	1656	4	US-09-489-039A-1878	Sequence 3878, Ap
117	-17	0.3	515	4	US-09-621-976-3443	Sequence 16, Appl	C 190	17	0.3	1669	4	US-09-591-466C-1	Sequence 1, Appl
118	-17	0.3	515	4	US-09-085-199B-16	Sequence 102, App	C 191	17	0.3	1671	4	US-09-795-691-3	Sequence 3, Appl
119	-17	0.3	516	3	US-08-691-814B-102	Sequence 923, App	C 192	17	0.3	1737	4	US-09-591-466C-3	Sequence 3, Appl
120	-17	0.3	520	2	US-08-986-171E-323	Sequence 1059, Ap	C 193	17	0.3	1747	1	US-08-176-620A-1	Sequence 1, Appl
121	-17	0.3	554	4	US-09-621-976-1059	Sequence 84, Appl	C 194	17	0.3	1747	1	US-08-463-862-1	Sequence 1, Appl
122	-17	0.3	567	4	US-09-205-258-84	Sequence 3321, Ap	C 195	17	0.3	1747	2	US-08-461-985-1	Sequence 1, Appl
123	-17	0.3	573	4	US-09-252-991A-3321	Sequence 2, Appl	C 196	17	0.3	1747	2	US-08-458-887-1	Sequence 1, Appl
124	-17	0.3	573	4	US-08-801-742-2	Sequence 50, Appl	C 197	17	0.3	1747	3	US-08-532-787B-1	Sequence 1, Appl
125	-17	0.3	577	4	US-09-340-620A-50	Sequence 2, Appl	C 198	17	0.3	1747	3	US-08-532-012C-1	Sequence 1, Appl
126	-17	0.3	585	4	US-08-808-303-1	Sequence 1, Appl	C 199	17	0.3	1872	4	US-09-280-116-110	Sequence 110, Appl
127	-17	0.3	588	1	US-08-808-303-2	Sequence 2, Appl	C 200	17	0.3	1872	4	US-08-888-818C-1	Sequence 1, Appl
128	-17	0.3	588	1	US-08-808-303-2	Sequence 1, Appl	C 201	17	0.3	1875	4	US-09-252-991A-5054	Sequence 5054, Ap
129	-17	0.3	588	3	US-08-996-533-1	Sequence 4, Appl	C 202	17	0.3	1904	1	US-08-412-375A-5	Sequence 5, Appl
130	-17	0.3	588	3	US-08-996-533-2	Sequence 1745, Ap	C 203	17	0.3	1926	3	US-09-117-250-4	Sequence 4, Appl
131	-17	0.3	600	3	US-09-328-111-433	Sequence 433, App	C 204	17	0.3	1962	4	US-09-352-991A-520	Sequence 520, Ap
132	-17	0.3	627	3	US-09-107-532A-1745	Sequence 359, App	C 205	17	0.3	1962	4	US-09-620-312D-715	Sequence 715, App
133	-17	0.3	656	3	US-08-998-416-56	Sequence 55, App	C 206	17	0.3	1990	1	US-08-454-720C-35	Sequence 932, App
134	-17	0.3	694	4	US-09-976-594-359	Sequence 686, App	C 207	17	0.3	1999	4	US-09-620-312D-932	Sequence 4, Appl
135	-17	0.3	697	3	US-09-328-111-686	Sequence 7, Appl	C 208	17	0.3	1999	4	US-09-167-109-4	Sequence 12, Appl
136	-17	0.3	736	1	US-08-808-303-7	Sequence 48, Appl	C 209	17	0.3	2003	4	US-09-921-099A-12	Sequence 1, Appl
137	-17	0.3	736	1	US-08-996-533-7	Sequence 1308, Ap	C 210	17	0.3	2003	4	US-08-691-814B-231	Sequence 231, App
138	-17	0.3	740	4	US-09-340-620A-48	Sequence 1601, Ap	C 211	17	0.3	2063	3	US-08-724-586-1	Sequence 1, Appl
139	-17	0.3	758	4	US-09-833-381-1308	Sequence 9665, Ap	C 212	17	0.3	2066	3	US-09-421-632-1	Sequence 1, Appl
140	-17	0.3	774	4	US-09-252-991A-1601	Sequence 595, App	C 213	17	0.3	2086	4	US-09-312-190-1	Sequence 1, Appl
141	-17	0.3	774	4	US-09-623-655-595	Sequence 1653, Ap	C 214	17	0.3	2091	1	US-09-813-818-1	Sequence 1, Appl
142	-17	0.3	779	4	US-09-452-991A-1652	Sequence 2658, Ap	C 215	17	0.3	2117	2	US-08-655-640-3	Sequence 3, Appl
143	-17	0.3	780	4	US-09-489-039A-2658	Sequence 132, App	C 216	17	0.3	2145	2	US-09-366-257-1	Sequence 1, Appl
144	-17	0.3	791	4	US-08-936-165A-132	Sequence 1865, Ap	C 217	17	0.3	2165	4	US-09-220-132-61	Sequence 61, Appl
145	-17	0.3	804	4	US-09-489-039A-1865	Sequence 5034, Ap	C 218	17	0.3	2177	4	US-09-016-834-1084	Sequence 1084, Ap
146	-17	0.3	828	4	US-09-352-991A-3247	Sequence 5573, Ap	C 219	17	0.3	2229	4	US-09-252-991A-7718	Sequence 5162, Ap
147	-17	0.3	854	4	US-09-489-039A-5573	Sequence 258, App	C 220	17	0.3	2229	4	US-08-154-019-1	Sequence 1, Appl
148	-17	0.3	854	4	US-09-489-039A-5573	Sequence 258, App	C 221	17	0.3	2319	1	US-08-461-333-1	Sequence 1, Appl
149	-17	0.3	948	4	US-09-641-638-258	Sequence 64, Appl	C 222	17	0.3	2319	1	US-08-464-167-1	Sequence 1, Appl
150	-17	0.3	1001	4	US-09-641-638-258	Sequence 96, App	C 223	17	0.3	2319	1	US-08-464-167-1	Sequence 1, Appl
151	-17	0.3	1038	4	US-09-665-189A-64	Sequence 496, App	C 224	17	0.3	2319	1	US-08-464-167-1	Sequence 1, Appl
152	-17	0.3	1039	4	US-09-221-017B-496	Sequence 10, Appl	C 225	17	0.3	2319	1	US-08-464-167-1	Sequence 1, Appl
153	-17	0.3	1046	4	US-09-716-161A-3	Sequence 21, Appl	C 226	17	0.3	2319	1	US-08-464-167-1	Sequence 1, Appl
154	-17	0.3	1096	4	US-09-716-161A-10	Sequence 26, Appl	C 227	17	0.3	2319	1	US-08-464-167-1	Sequence 1, Appl
155	-17	0.3	1112	4	US-09-543-681A-432	Sequence 26, Appl	C 228	17	0.3	2319	1	US-08-464-167-1	Sequence 1, Appl
156	-17	0.3	1137	4	US-09-491-577-21	Sequence 5101, Ap	C 229	17	0.3	2360	1	US-08-453-703-1	Sequence 1, Appl
157	-17	0.3	1155	4	US-08-107-676-26	Sequence 5101, Ap	C 230	17	0.3	2360	1	US-08-453-703-1	Sequence 1, Appl
158	-17	0.3	1178	4	US-09-295-820-26	Sequence 2, Appl	C 231	17	0.3	2360	1	US-08-453-703-1	Sequence 1, Appl
159	-17	0.3	1206	4	US-09-252-991A-5101	Sequence 2, Appl	C 232	17	0.3	2360	1	US-08-453-703-1	Sequence 1, Appl
160	-17	0.3	1211	2	US-08-107-676-2	Sequence 1, Appl	C 233	17	0.3	2360	1	US-08-453-703-1	Sequence 1, Appl
161	-17	0.3	1211	2	US-08-107-676-2	Sequence 1, Appl	C 234	17	0.3	2360	1	US-08-453-703-1	Sequence 1, Appl
162	-17	0.3	1212	1	US-08-282-581-1	Sequence 1, Appl	C 235	17	0.3	2360	1	US-08-453-703-1	Sequence 1, Appl
163	-17	0.3	1212	1	US-08-282-581-1	Sequence 1, Appl	C 236	17	0.3	2360	1	US-08-453-703-1	Sequence 1, Appl
164	-17	0.3	1212	1	US-08-282-581-1	Sequence 1, Appl	C 237	17	0.3	2360	1	US-08-453-703-1	Sequence 1, Appl
165	-17	0.3	1212	1	US-08-282-581-1	Sequence 1, Appl	C 238	17	0.3	2360	1	US-08-453-703-1	Sequence 1, Appl
166	-17	0.3	1212	1	US-08-282-581-1	Sequence 1, Appl	C 239	17	0.3	2360	1	US-08-453-703-1	Sequence 1, Appl
167	-17	0.3	1212	1	US-08-282-581-1	Sequence 1, Appl	C 240	17	0.3	2360	1	US-08-453-703-1	Sequence 1, Appl
168	-17	0.3	1212	1	US-08-282-581-1	Sequence 1, Appl	C 241	17	0.3	2360	1	US-08-453-703-1	Sequence 1, Appl
169	-17	0.3	1212	1	US-08-282-581-1	Sequence 1, Appl	C 242	17	0.3	2360	1	US-08-453-703-1	Sequence 1, Appl
170	-17	0.3	1212	1	US-08-282-581-1	Sequence 1, Appl	C 243	17	0.3	2360	1	US-08-453-703-1	Sequence 1, Appl
171	-17	0.3	1212	1	US-08-282-581-1	Sequence 1, Appl	C 244	17	0.3	2360	1	US-08-453-703-1	Sequence 1, Appl
172	-17	0.3	1212	1	US-08-282-581-1	Sequence 1, Appl	C 245	17	0.3	2360	1	US-08-453-703-1	Sequence 1, Appl
173	-17	0.3	1212	1	US-08-282-581-1	Sequence 1, Appl	C 246	17	0.3	2360	1	US-08-453-703-1	Sequence 1, Appl

247	17	0.3	2619	4	US-09-377-497-14	Sequence 14, Appl	320	16	0.3	131	3	US-09-099-639-62	Sequence 62, Appl
248	17	0.3	2686	4	US-09-795-691-1	Sequence 1, Appl	321	16	0.3	131	5	PCT-US93-12588-62	Sequence 62, Appl
C 249	17	0.3	3015	4	US-09-489-039A-2544	Sequence 2544, Ap	322	16	0.3	119	5	PCT-US95-08071-62	Sequence 62, Appl
C 250	17	0.3	3024	2	US-08-836-943-1	Sequence 1, Appl	C 323	16	0.3	157	4	US-09-621-976-15276	Sequence 15276, A
C 251	17	0.3	3034	4	US-09-976-594-636	Sequence 636, App	C 324	16	0.3	157	4	US-09-621-976-15275	Sequence 15275, A
C 252	17	0.3	3060	4	US-09-252-991A-7846	Sequence 7846, Ap	C 325	16	0.3	157	4	US-09-621-976-15278	Sequence 15278, A
253	17	0.3	3366	4	US-09-345-650-2	Sequence 2, Appl	326	16	0.3	182	4	US-09-313-294A-4709	Sequence 4709, Ap
254	17	0.3	3458	4	US-09-923-655-603	Sequence 603, App	327	16	0.3	186	4	US-09-705-448-5	Sequence 5, Appl
255	17	0.3	3505	1	US-08-454-097-34	Sequence 34, Appl	C 328	16	0.3	210	4	US-09-134-001C-2749	Sequence 2749, Ap
256	17	0.3	3505	1	US-08-468-036-43	Sequence 43, Appl	C 329	16	0.3	225	4	US-09-621-976-15277	Sequence 15277, A
257	17	0.3	3505	2	US-08-376-843-43	Sequence 43, Appl	C 330	16	0.3	224	4	US-09-621-976-15277	Sequence 15277, A
258	17	0.3	3505	3	US-08-185-359-34	Sequence 34, Appl	C 331	16	0.3	226	4	US-09-621-976-15272	Sequence 15272, A
259	17	0.3	3601	3	US-09-017-631-23	Sequence 23, Appl	332	16	0.3	227	4	US-09-705-448-6	Sequence 6, Appl
260	17	0.3	3602	2	US-08-683-795A-33	Sequence 33, Appl	333	16	0.3	229	4	US-09-833-381-18	Sequence 18, Appl
261	17	0.3	3602	3	US-09-018-138-1	Sequence 1, Appl	C 334	16	0.3	247	4	US-09-016-434-436	Sequence 436, App
262	17	0.3	3900	4	US-09-023-655-1440	Sequence 1420, Ap	C 335	16	0.3	248	4	US-09-016-434-436	Sequence 968, App
263	17	0.3	3984	4	US-09-016-434-1199	Sequence 1199, Ap	C 336	16	0.3	258	4	US-09-282-991A-7336	Sequence 7356, Ap
264	17	0.3	4307	4	US-09-552-950-2	Sequence 2, Appl	337	16	0.3	270	4	US-09-282-991A-6796	Sequence 5174, Ap
265	17	0.3	4348	2	US-08-915-868-1	Sequence 1, Appl	338	16	0.3	280	4	US-09-313-294A-6796	Sequence 6796, Ap
C 266	17	0.3	4530	4	US-09-566-921-134	Sequence 134, Appl	339	16	0.3	283	2	US-09-313-294A-7265	Sequence 7265, Ap
C 267	17	0.3	4530	4	US-09-566-921-134	Sequence 61, Appl	340	16	0.3	289	4	US-09-382-552-56	Sequence 8492, Ap
268	17	0.3	4633	4	US-09-453-702B-22	Sequence 22, Appl	341	16	0.3	292	4	US-09-621-976-8492	Sequence 8492, Ap
269	17	0.3	4632	4	US-09-421-017B-846	Sequence 846, App	342	16	0.3	315	4	US-09-313-294A-5804	Sequence 5804, Ap
270	17	0.3	5124	4	US-09-534-638-2	Sequence 964, App	343	16	0.3	354	4	US-09-489-039A-2664	Sequence 2664, Ap
271	17	0.3	5129	4	US-09-976-594-964	Sequence 1, Appl	C 344	16	0.3	360	4	US-09-621-976-7422	Sequence 2422, Ap
272	17	0.3	5330	4	US-09-023-905A-1	Sequence 5, Appl	C 345	16	0.3	368	4	US-09-377-787-743	Sequence 143, App
273	17	0.3	5889	1	US-08-463-092B-5	Sequence 5, Appl	346	16	0.3	384	4	US-09-134-000C-1544	Sequence 1544, App
274	17	0.3	5889	2	US-08-462-109A-5	Sequence 5, Appl	347	16	0.3	396	4	US-09-640-173-147	Sequence 147, App
275	17	0.3	5889	3	US-08-463-107B-5	Sequence 5, Appl	C 348	16	0.3	411	4	US-09-713-550-147	Sequence 712, App
276	17	0.3	5889	3	US-08-463-179A-5	Sequence 5, Appl	C 350	16	0.3	414	4	US-09-489-039A-712	Sequence 11049, A
277	17	0.3	5889	3	US-08-461-384B-5	Sequence 15, Appl	C 351	16	0.3	420	3	US-09-621-976-11049	Sequence 5, Appl
278	17	0.3	6002	1	US-08-698-551-15	Sequence 15, Appl	C 352	16	0.3	420	3	US-09-532-803-E	Sequence 766, App
279	17	0.3	6002	2	US-08-602-228-15	Sequence 15, Appl	C 353	16	0.3	420	3	US-09-252-991A-756	Sequence 12, Appl
280	17	0.3	6002	2	US-08-839-032A-15	Sequence 15, Appl	C 354	16	0.3	421	3	US-09-211-631-12	Sequence 12, Appl
281	17	0.3	6002	2	US-08-839-032A-15	Sequence 15, Appl	C 355	16	0.3	421	3	US-09-265-628-12	Sequence 10, Appl
282	17	0.3	6002	2	US-08-839-032A-15	Sequence 15, Appl	C 356	16	0.3	421	3	US-09-001-141-10	Sequence 13, Appl
C 283	17	0.3	6827	4	US-08-961-527-60	Sequence 60, Appl	C 357	16	0.3	421	4	US-09-653-403-13	Sequence 321, Appl
C 284	17	0.3	6976	4	US-09-636-215-705	Sequence 705, App	C 358	16	0.3	421	4	US-09-644-597-321	Sequence 13, Appl
C 285	17	0.3	7063	4	US-09-685-166A-705	Sequence 5, Appl	C 359	16	0.3	421	4	US-10-013-784-13	Sequence 321, App
C 286	17	0.3	7812	4	US-09-485-632B-5	Sequence 9, Appl	C 360	16	0.3	421	4	US-09-480-884A-321	Sequence 321, App
C 287	17	0.3	8266	4	US-09-485-632B-9	Sequence 54, Appl	C 361	16	0.3	421	4	US-09-542-615A-321	Sequence 321, App
288	17	0.3	8266	4	US-09-485-632B-9	Sequence 54, Appl	C 362	16	0.3	421	4	US-09-606-421B-321	Sequence 11, Appl
289	17	0.3	9772	4	US-09-552-950-5	Sequence 3, Appl	C 363	16	0.3	421	4	US-09-786-606-11	Sequence 8508, Ap
C 290	17	0.3	14707	1	US-09-312-762A-3	Sequence 1, Appl	C 364	16	0.3	423	4	US-09-621-976-6508	Sequence 11873, A
C 291	17	0.3	30001	2	US-08-125-468-1	Sequence 1, Appl	C 365	16	0.3	423	4	US-09-252-991A-11873	Sequence 382, App
C 292	17	0.3	34001	2	US-08-474-933-1	Sequence 18, Appl	C 366	16	0.3	423	4	US-09-489-039A-382	Sequence 1140, Ap
C 293	17	0.3	47981	4	US-09-679-279-1	Sequence 1, Appl	C 367	16	0.3	423	4	US-09-976-594-765	Sequence 18939, A
C 294	17	0.3	58909	4	US-09-596-002-30	Sequence 30, Appl	C 368	16	0.3	429	4	US-09-976-594-765	Sequence 18939, A
295	17	0.3	75395	4	US-09-984-890-3	Sequence 3, Appl	369	16	0.3	431	4	US-09-621-976-18939	Sequence 18230, A
296	17	0.3	75395	4	US-09-410-551B-1	Sequence 10, Appl	370	16	0.3	437	4	US-09-621-976-18230	Sequence 15600, A
297	17	0.3	98844	4	US-09-791-211-70	Sequence 32, Appl	371	16	0.3	446	4	US-09-666-751-59	Sequence 59, Appl
298	17	0.3	118067	4	US-09-497-855A-32	Sequence 40, Appl	372	16	0.3	450	4	US-09-621-976-18938	Sequence 18938, A
C 299	17	0.3	161652	4	US-09-497-855A-40	Sequence 24, Appl	C 373	16	0.3	451	4	US-09-621-976-18938	Sequence 17249, A
C 300	17	0.3	169998	4	US-09-676-610B-24	Sequence 17, Appl	C 374	16	0.3	457	4	US-09-621-976-17249	Sequence 3878, Ap
C 301	17	0.3	176373	3	US-09-128-155-17	Sequence 10, Appl	C 375	16	0.3	470	4	US-09-621-976-1838	Sequence 174, App
C 302	17	0.3	197496	4	US-09-877-177A-10	Sequence 1, Appl	C 376	16	0.3	478	4	US-09-280-116-174	Sequence 893, App
C 303	17	0.3	580073	4	US-08-545-588D-1	Sequence 3, Appl	C 377	16	0.3	478	4	US-09-621-976-893	Sequence 1337, Ap
C 304	17	0.3	786431	4	US-09-751-389-3	Sequence 16, Appl	C 378	16	0.3	480	4	US-09-621-976-1537	Sequence 564, App
C 305	16	0.3	24	4	US-09-599-287A-16	Sequence 16, Appl	C 379	16	0.3	481	4	US-09-081-072-9	Sequence 9, Appl
C 306	16	0.3	25	4	US-09-363-939A-164	Sequence 168, App	C 380	16	0.3	489	1	US-08-444-093A-9	Sequence 565, App
C 307	16	0.3	65	4	US-09-363-939A-168	Sequence 36, Appl	C 381	16	0.3	492	1	US-09-621-976-565	Sequence 153, App
C 308	16	0.3	78	3	US-08-479-744A-36	Sequence 36, Appl	C 382	16	0.3	494	4	US-09-280-116-153	Sequence 3245, Ap
C 309	16	0.3	78	3	US-08-280-757B-36	Sequence 5, Appl	C 383	16	0.3	498	4	US-09-489-039A-1543	Sequence 1541, Ap
310	16	0.3	78	3	US-09-227-555-5	Sequence 3, Appl	C 384	16	0.3	501	4	US-09-328-352-3345	Sequence 8727, Ap
311	16	0.3	78	3	US-09-425-762-36	Sequence 5, Appl	C 385	16	0.3	510	4	US-09-489-039A-1541	Sequence 8727, Ap
312	16	0.3	80	3	US-09-133-321-9	Sequence 88, Appl	C 386	16	0.3	534	4	US-09-252-991A-4885	Sequence 14895, A
313	16	0.3	129	3	US-09-209-676-88	Sequence 62, Appl	C 387	16	0.3	540	3	US-09-331-625A-44	Sequence 44, Appl
314	16	0.3	131	1	US-08-453-274B-62	Sequence 62, Appl	C 388	16	0.3	546	3	US-09-494-151-44	Sequence 44, Appl
315	16	0.3	131	1	US-08-453-274B-62	Sequence 62, Appl	C 389	16	0.3	546	3	US-09-494-151-44	Sequence 44, Appl
316	16	0.3	131	1	US-08-453-274B-62	Sequence 62, Appl	C 390	16	0.3	546	3	US-09-494-151-44	Sequence 44, Appl
317	16	0.3	131	1	US-08-453-274B-62	Sequence 62, Appl	C 391	16	0.3	546	3	US-09-494-151-44	Sequence 44, Appl
318	16	0.3	131	1	US-08-453-274B-62	Sequence 62, Appl	C 392	16	0.3	546	3	US-09-494-151-44	Sequence 44, Appl
319	16	0.3	131	2	US-08-453-274B-62	Sequence 62, Appl	C 392	16	0.3	546	3	US-09-494-151-44	Sequence 44, Appl

C 393	16	0.3	552	4	US-09-489-039A-3494	Sequence 3494, Ap	C 466	16	0.3	954	4	US-10-043-238-2	Sequence 2, Appl1
C 394	16	0.3	571	4	US-09-404-879A-48	Sequence 48, Appl	C 467	16	0.3	963	4	US-09-071-038-285	Sequence 285, App
C 395	16	0.3	571	4	US-09-338-933-48	Sequence 48, Appl	C 468	16	0.3	963	4	US-09-134-000C-4132	Sequence 2133, Ap
C 396	16	0.3	571	4	US-09-215-681-48	Sequence 48, Appl	C 469	16	0.3	966	4	US-09-252-991A-1439	Sequence 7433, Ap
C 397	16	0.3	571	4	US-09-216-003A-48	Sequence 48, Appl	C 470	16	0.3	972	4	US-09-252-991A-1956	Sequence 2956, Ap
C 398	16	0.3	576	1	US-08-086-428B-2	Sequence 2, Appl1	C 471	16	0.3	987	4	US-09-489-039A-1556	Sequence 2674, Ap
C 399	16	0.3	576	1	US-08-468-570-2	Sequence 2, Appl1	C 472	16	0.3	990	4	US-09-252-991A-11517	Sequence 11517, A
C 400	16	0.3	576	2	US-08-290-665A-2	Sequence 2, Appl1	C 473	16	0.3	1001	4	US-09-641-638-270	Sequence 270, App
C 401	16	0.3	576	4	US-08-466-601A-2	Sequence 2, Appl1	C 474	16	0.3	1001	4	US-09-641-638-446	Sequence 446, App
C 402	16	0.3	576	4	PCT-US95-10398-2	Sequence 2, Appl1	C 475	16	0.3	1001	4	US-09-671-317-290	Sequence 290, App
C 403	16	0.3	579	4	US-09-252-991A-8637	Sequence 8637, Ap	C 476	16	0.3	1007	4	US-09-328-476C-106	Sequence 106, App
C 404	16	0.3	582	4	US-09-621-976-18934	Sequence 18934, A	C 477	16	0.3	1011	4	US-09-641-808-3	Sequence 3, Appl1
C 405	16	0.3	603	3	US-09-385-988-233	Sequence 233, App	C 478	16	0.3	1011	4	US-09-723-123-3	Sequence 3, Appl1
C 406	16	0.3	605	4	US-09-643-597-207	Sequence 207, App	C 479	16	0.3	1011	4	US-09-252-991A-14782	Sequence 14782, A
C 407	16	0.3	605	4	US-09-480-884A-207	Sequence 207, App	C 480	16	0.3	1011	4	US-09-722-862-3	Sequence 3, Appl1
C 408	16	0.3	605	4	US-09-543-615A-207	Sequence 207, App	C 481	16	0.3	1026	4	US-09-252-991A-6229	Sequence 6229, Ap
C 409	16	0.3	605	4	US-09-608-421B-207	Sequence 207, App	C 482	16	0.3	1026	4	US-09-641-808-1	Sequence 1, Appl1
C 410	16	0.3	615	1	US-08-539-304A-1	Sequence 1, Appl1	C 483	16	0.3	1026	4	US-09-723-123-1	Sequence 1, Appl1
C 411	16	0.3	615	1	US-09-134-001C-7663	Sequence 2763, Ap	C 484	16	0.3	1026	4	US-09-722-862-1	Sequence 1, Appl1
C 412	16	0.3	630	4	US-09-252-991A-656	Sequence 656, App	C 485	16	0.3	1026	4	US-09-252-991A-15460	Sequence 15460, A
C 413	16	0.3	631	3	US-08-545-809A-31	Sequence 31, Appl	C 486	16	0.3	1041	4	US-09-489-039A-194	Sequence 194, Appl
C 414	16	0.3	631	3	US-08-943-731-187	Sequence 187, App	C 487	16	0.3	1044	4	US-07-975-528-1	Sequence 1, Appl1
C 415	16	0.3	651	4	US-09-489-039A-957	Sequence 957, App	C 488	16	0.3	1044	4	US-07-974-409C-422	Sequence 422, App
C 416	16	0.3	654	4	US-09-252-991A-11814	Sequence 11814, A	C 489	16	0.3	1044	4	US-07-974-409C-423	Sequence 423, App
C 417	16	0.3	659	4	US-09-221-017B-61	Sequence 61, Appl	C 490	16	0.3	1047	4	US-09-252-991A-1034	Sequence 9034, Ap
C 418	16	0.3	672	4	US-09-489-039A-6264	Sequence 6264, Ap	C 491	16	0.3	1065	3	US-09-325-897-1	Sequence 1, Appl1
C 419	16	0.3	688	3	US-09-328-111-453	Sequence 453, App	C 492	16	0.3	1065	4	US-09-837-728-1	Sequence 1, Appl1
C 420	16	0.3	696	1	US-07-656-566-4	Sequence 4, Appl1	C 493	16	0.3	1071	4	US-09-543-681A-3985	Sequence 3985, Ap
C 421	16	0.3	697	6	5171843-10	Patent No. 5171843	C 494	16	0.3	1086	3	US-09-280-409-3	Sequence 3, Appl1
C 422	16	0.3	701	3	US-08-331-625A-55	Sequence 55, Appl	C 495	16	0.3	1088	4	US-09-755-665-61	Sequence 61, Appl
C 423	16	0.3	701	3	US-09-133-321-1	Sequence 1, Appl1	C 496	16	0.3	1090	4	US-09-755-665-59	Sequence 59, Appl
C 424	16	0.3	701	4	US-09-494-151-55	Sequence 55, Appl	C 497	16	0.3	1090	4	US-09-755-665-60	Sequence 60, Appl
C 425	16	0.3	701	4	US-09-972-484-55	Sequence 55, Appl	C 498	16	0.3	1091	4	US-09-755-665-62	Sequence 62, Appl
C 426	16	0.3	701	4	US-09-621-976-206	Sequence 206, App	C 499	16	0.3	1098	4	US-09-252-991A-15400	Sequence 15400, A
C 427	16	0.3	703	4	US-09-023-655-1097	Sequence 1097, App	C 500	16	0.3	1101	4	US-09-424-712-1	Sequence 1, Appl1
C 428	16	0.3	710	3	US-08-998-416-603	Sequence 603, App	C 501	16	0.3	1103	4	US-09-424-712-2	Sequence 2, Appl1
C 429	16	0.3	715	4	US-09-247-155-139	Sequence 139, App	C 502	16	0.3	1110	4	US-09-543-681A-3342	Sequence 3342, Ap
C 430	16	0.3	726	4	US-09-621-976-1875	Sequence 1875, Ap	C 503	16	0.3	1123	4	US-09-755-665-15	Sequence 15, Appl
C 431	16	0.3	729	4	US-09-252-991A-11877	Sequence 11877, A	C 504	16	0.3	1137	6	5171843-8	Patent No. 5171843
C 432	16	0.3	730	2	US-09-352-990-3	Sequence 3, Appl1	C 505	16	0.3	1139	4	US-09-690-454-13	Sequence 13, Appl
C 433	16	0.3	738	2	US-08-604-989A-8	Sequence 8, Appl1	C 506	16	0.3	1140	4	US-09-252-991A-14999	Sequence 14999, A
C 434	16	0.3	741	4	US-09-149-476-202	Sequence 202, App	C 507	16	0.3	1140	4	US-09-328-352-4089	Sequence 4089, Ap
C 435	16	0.3	741	4	US-09-252-991A-12066	Sequence 12066, A	C 508	16	0.3	1149	4	US-09-252-991A-14783	Sequence 14783, A
C 436	16	0.3	741	4	US-09-489-039A-5512	Sequence 5512, Ap	C 509	16	0.3	1176	4	US-09-543-681A-3917	Sequence 3917, Ap
C 437	16	0.3	742	2	US-08-966-316-2	Sequence 2, Appl1	C 510	16	0.3	1176	4	US-09-489-039A-5553	Sequence 5553, Ap
C 438	16	0.3	765	3	US-08-331-625A-40	Sequence 40, Appl	C 511	16	0.3	1179	2	US-08-465-794-4	Sequence 4, Appl1
C 439	16	0.3	765	4	US-09-494-151-40	Sequence 40, Appl	C 512	16	0.3	1179	3	US-09-049-878-4	Sequence 4, Appl1
C 440	16	0.3	765	4	US-09-252-991A-7969	Sequence 7969, Ap	C 513	16	0.3	1182	1	US-07-640-476-4	Sequence 4, Appl1
C 441	16	0.3	765	4	US-09-972-484-40	Sequence 40, Appl	C 514	16	0.3	1182	4	US-09-252-991A-6534	Sequence 6534, Ap
C 442	16	0.3	768	4	US-09-252-991A-9221	Sequence 9221, Ap	C 515	16	0.3	1185	6	5290690-1	Patent No. 5290690
C 443	16	0.3	780	4	US-09-651-656-26	Sequence 26, Appl	C 516	16	0.3	1194	4	US-09-252-991A-10197	Sequence 10197, A
C 444	16	0.3	780	4	US-09-650-855-26	Sequence 26, Appl	C 517	16	0.3	1197	4	US-09-891-641-43	Sequence 43, Appl
C 445	16	0.3	783	4	US-09-252-991A-11569	Sequence 11569, A	C 518	16	0.3	1203	4	US-09-540-236-1833	Sequence 823, App
C 446	16	0.3	793	4	US-09-252-991A-3048	Sequence 3048, Ap	C 519	16	0.3	1203	4	US-09-489-039A-823	Sequence 823, App
C 447	16	0.3	813	4	US-09-252-991A-12621	Sequence 12621, A	C 520	16	0.3	1212	4	US-09-252-991A-12510	Sequence 12510, A
C 448	16	0.3	826	3	US-09-071-035-287	Sequence 287, App	C 521	16	0.3	1215	4	US-09-543-681A-3574	Sequence 3574, Ap
C 449	16	0.3	835	3	US-08-957-302A-9	Sequence 9, Appl1	C 522	16	0.3	1215	4	US-09-543-681A-3574	Sequence 3574, Ap
C 450	16	0.3	835	3	US-09-542-403-9	Sequence 9, Appl1	C 523	16	0.3	1226	4	US-09-328-352-1841	Sequence 1841, Ap
C 451	16	0.3	836	3	US-08-998-416-332	Sequence 332, Appl	C 524	16	0.3	1226	4	US-09-540-236-1436	Sequence 1436, Ap
C 452	16	0.3	840	4	US-09-252-991A-10769	Sequence 10769, A	C 525	16	0.3	1245	4	US-09-252-991A-14749	Sequence 14749, A
C 453	16	0.3	840	4	US-09-252-991A-14967	Sequence 14967, A	C 526	16	0.3	1245	1	US-07-887-078B-1	Sequence 1, Appl1
C 454	16	0.3	840	4	US-09-252-991A-14967	Sequence 14967, A	C 527	16	0.3	1245	1	US-08-466-444-1	Sequence 1, Appl1
C 455	16	0.3	843	4	US-09-252-991A-2744	Sequence 2744, Ap	C 528	16	0.3	1248	4	US-09-252-991A-6318	Sequence 6318, Ap
C 456	16	0.3	858	4	US-09-172-965-8	Sequence 8, Appl1	C 529	16	0.3	1248	4	US-09-252-991A-9872	Sequence 9872, Ap
C 457	16	0.3	870	4	US-09-489-039A-764	Sequence 764, App	C 530	16	0.3	1251	4	US-09-540-236-295	Sequence 295, Appl
C 458	16	0.3	876	4	US-09-513-783A-9	Sequence 9, Appl1	C 531	16	0.3	1252	4	US-09-673-395A-73	Sequence 73, Appl
C 459	16	0.3	876	4	US-09-252-991A-8946	Sequence 8946, Ap	C 532	16	0.3	1257	4	US-09-489-039A-6981	Sequence 6981, Ap
C 460	16	0.3	876	4	US-09-252-991A-14476	Sequence 14476, A	C 533	16	0.3	1284	3	US-08-331-625A-42	Sequence 42, Appl
C 461	16	0.3	891	4	US-09-489-039A-360	Sequence 360, App	C 534	16	0.3	1284	4	US-09-692-484-42	Sequence 42, Appl
C 462	16	0.3	891	4	PCT-US96-0530A-303	Sequence 303, App	C 535	16	0.3	1284	4	US-09-972-484-42	Sequence 42, Appl
C 463	16	0.3	915	3	US-08-557-614-1	Sequence 1, Appl1	C 536	16	0.3	1287	4	US-09-489-039A-394	Sequence 394, App
C 464	16	0.3	918	4	US-09-252-991A-12009	Sequence 12009, A	C 537	16	0.3	1299	1	US-08-723-938-4	Sequence 4, Appl1
C 465	16	0.3	933	4	US-09-252-991A-8606	Sequence 8606, Ap	C 538	16	0.3	1299	2	US-09-080-538-4	Sequence 4, Appl1